

## THESIS / THÈSE

### MASTER EN SCIENCES INFORMATIQUES

#### Aide au diagnostic médical

#### conception et analyse d'un système d'information à l'usage du centre anti-poisons de Bruxelles

Biernaux, Thierry

*Award date:*  
1979

*Awarding institution:*  
Université de Namur

[Link to publication](#)

#### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

#### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

FACULTES UNIVERSITAIRES NOTRE-DAME DE LA PAIX  
A NAMUR  
INSTITUT D'INFORMATIQUE

Année Académique 1978-1979

# **AIDE AU DIAGNOSTIC MEDICAL**

**CONCEPTION ET ANALYSE D'UN SYSTEME D'INFORMATION**

**A L'USAGE DU CENTRE ANTI-POISONS DE BRUXELLES**

**THIERRY BIERNAUX**

Mémoire présenté en vue de l'obtention  
du grade de licencié et maître en Informatique.

**DIRECTEUR : MR J. FICHEFET**



FACULTES  
UNIVERSITAIRES  
N.-D. DE LA PAIX  
NAMUR

Bibliothèque

FMB 16/

1979/20

FMB 16 / 1979 / 20

## REMERCIEMENTS.

Nous tenons à exprimer une profonde gratitude à toutes les personnes qui nous ont aidés à mener notre travail à bien, par leur direction, leurs conseils et leur aide :

Monsieur J. Fichet (directeur du mémoire),  
Madame R. Gerday (directrice des Banques de Données Toxicologiques de l'INSERM, Paris),  
Madame A. Govaerts (directrice du Centre Anti-Poisons de Bruxelles).

Qu'il nous soit également permis de remercier  
Messieurs M. Le Minor, J-P. Nakache et B. Asselain (équipes de recherches de l'INSERM) et  
Madame Riboulet (Centre Anti-Poisons de Paris),  
qui nous ont permis de clarifier les difficultés rencontrées lors de la réalisation de cette étude.

FACULTES UNIVERSITAIRES NOTRE-DAME DE LA PAIX  
A NAMUR  
INSTITUT D'INFORMATIQUE

Année Académique 1978-1979

# **AIDE AU DIAGNOSTIC MEDICAL**

CONCEPTION ET ANALYSE D'UN SYSTEME D'INFORMATION

A L'USAGE DU CENTRE ANTI-POISONS DE BRUXELLES

**THIERRY BIERNAUX**

Mémoire présenté en vue de l'obtention  
du grade de licencié et maître en Informatique.

**DIRECTEUR : MR J. FICHEFET**

LBS 3620737  
73158



## TABLE DES MATIERES.

Introduction générale .....	1.
Ière partie : <u>L'aide au diagnostic médical. Méthodes</u> <u>et applications.</u>	
Introduction .....	3.
Chapitre 1 : <u>Informatique et médecine.</u>	
A. Prise de conscience du problème : une ère d'évolution .....	5.
B. Les domaines actuels d'application .....	5.
C. Les difficultés de mise en place .....	6.
D. Problèmes de fond : le secret médical, le conflit de compétence et la responsabilité .....	7.
§1. Le secret médical .....	7.
§2. La compétence du médecin .....	9.
§3. La responsabilité .....	9.
Chapitre 2 : <u>L'aide au diagnostic médical.</u>	
A. Caractéristiques générales .....	11.
§1. Quelques définitions .....	11.
§2. La problématique du diagnostic médical .....	11.
§3. Shématisation de la démarche diagnostique .....	12.
B. Automatisation de l'aide au diagnostic .....	13.
§1. L'objectif .....	13.
§2. L'apport de l'informatique .....	13.
§3. Les limites de compétence .....	14.
Chapitre 3 : <u>Méthodologie.</u>	
A. Les méthodes d'aide au diagnostic .....	16.
Avertissement .....	16.
§1. Le diagnostic global .....	16.
1.1. Les méthodes d'analyse de données multidimen- sionnelle .....	16.
1.2. Les méthodes de discrimination non paramé- trique .....	22.
§2. Le diagnostic pas à pas .....	23.
2.1. Les méthodes logiques .....	23.
2.1.1. Application du calcul booléen .....	23.

2.1.2. Extension à l'algèbre ternaire .....	34.
2.1.3. La démonstration automatique de théorèmes ..	35.
2.2. La méthode probabiliste .....	36.
2.3. La théorie des sous-ensembles flous .....	41.
B. Les méthodes de choix d'examens complémentaires ..	43.
§1. Théorie de la décision .....	43.
§2. Théorie de l'information .....	45.
§3. Heuristiques .....	47.
3.1. Méthode S .....	47.
3.2. Méthode $\Sigma$ .....	47.
3.3. Méthode $\Sigma S$ .....	48.
3.4. Pas à pas .....	48.
Chapitre 4 : <u>Deux exemples d'application.</u>	
A. Application des méthodes d'analyse multidimension- nelle au pronostic de l'infarctus myocardique ....	49.
§1. Le problème thérapeutique .....	49.
§2. Les variables utilisées dans l'étude .....	49.
§3. Etude de l'échantillon de base par l'Analyse des Correspondances .....	50.
§4. Pronostic et décision thérapeutique .....	51.
§5. Conclusion .....	53.
B. Application de la méthode probabiliste au diagnostic en neurologie .....	53.
§1. Origine .....	53.
§2. Source des données .....	54.
§3. Saisie des données .....	54.
§4. L'outil informatique .....	54.
§5. Exploitation et utilisation .....	59.
§6. Analyse des résultats .....	59.
6.1. Analyse globale .....	59.
6.2. Analyse par diagnostic .....	61.
6.3. Apport pédagogique .....	63.
§7. Critique .....	64.
Chapitre 5 : <u>Informatique et Médecine. Bilan et perspectives.</u>	
A. La surveillance automatique des grands malades ...	66.
B. Les banques de données médicales .....	67.
C. L'aide au diagnostic .....	68.



## IIème partie : Application à la toxicologie clinique d'urgence.

Introduction .....	71.
Chapitre 1 : <u>La toxicologie. Fixation des limites de</u> <u>responsabilité. Rôle du C.A.P.</u>	
A. Les intoxications aiguës .....	72.
B. Les intoxications subaiguës .....	73.
C. Les intoxications chroniques .....	73.
Chapitre 2 : <u>Structure des C.A.P.</u>	
A. Pays-Bas .....	76.
B. Grande-Bretagne .....	77.
C. Suisse .....	77.
D. Belgique .....	77.
E. France .....	78.
F. Le centre 'idéal' .....	79.
Chapitre 3 : <u>L'impact de l'informatique au niveau</u> <u>des Centres Anti-Poisons.</u>	
A. Caractéristiques générales de l'information toxicologique .....	81.
B. Etude d'un choix informatique .....	82.
§1. Saisie et exploitation des dossiers d'appel ....	82.
§2. Automatisation du C.A.P. de Lyon .....	83.
Chapitre 4 : <u>Application au centre anti-poisons</u> <u>de Bruxelles.</u>	
A. Rappel des fonctions essentielles .....	84.
B. Etude des besoins .....	85.
C. Proposition d'éléments de réponse .....	87.
D. Vue globale du système proposé .....	88.
E. Etude détaillée du système proposé .....	89.
§1. Le catalogue des produits .....	89.
1.1. Impératifs du sous-système .....	89.
1.2. Les données .....	90.
1.2.1. Les produits .....	90.
1.2.2. Les informations relatives à ces produits .	91.
1.3. L'analyse des données .....	93.

1.3.A. Analyse logique des données .....	94.
1.3.A.1. Description des données individuelles ..	94.
1.3.A.2. Description des unités d'information ...	100.
1.3.A.3. Relations entre unités d'information ...	101.
1.3.A.4. Description des articles et des fichiers logiques .....	102.
1.3.A.5. Structure d'accès .....	105.
1.3.B. Organisation physique des données .....	106.
1.3.B.1. Définition physique des données individuelles .....	106.
1.3.B.2. Organisation physique des fichiers .....	111.
a. Le fichier direct .....	111.
b. Les dictionnaires de codification .....	112.
c. Le fichier inversé des noms-génériques et des noms-de-groupe .....	112.
d. Le fichier inversé des synonymes .....	113.
e. Le dictionnaire d'accès au catalogue .....	113.
e.1. Fixation des paramètres .....	114.
e.2. Calcul du nombre de niveaux nécessaires	115.
e.3. Représentation du dictionnaire d'accès	115.
f. Le fichier des annexes .....	117.
1.4. Analyse des traitements .....	119.
1.4.1. Création .....	119.
1.4.2. Mise à jour .....	122.
a. Ajoute .....	123.
b. Modification .....	123.
c. Suppression .....	125.
1.4.3. Consultation .....	125.
1.4.4. Edition .....	127.
1.4.5. Réorganisation .....	127.
1.4.6. Tri .....	128.
1.4.7. Recherches sur le catalogue .....	129.
§2. La gestion des dossiers d'appel .....	129.
2.1. Les banques de données de l'INSERM .....	130.
2.2. Le C.A.P. de Bruxelles. Etude et critique de l'existant .....	130.
2.3. Présentation du nouveau système .....	131.
2.3.1. Constitution de la fiche d'appel .....	131.
2.3.2. Contrôle direct .....	134.



2.3.3. Deuxième contrôle .....	134.
2.3.4. Saisie ou mise à jour des dossiers d'appel .....	134.
2.3.5. Archivage .....	135.
2.3.6. Consultation. Interrogation .....	135.
2.3.7. Complètement de la fiche d'appel .....	135.
2.3.8. Production d'états statistiques .....	135.
2.4. Analyse logique des données .....	136.
2.4.1. Description des données élémentaires .....	136.
2.4.2. Description des unités d'information .....	146.
2.4.3. Schéma conceptuel .....	147.
2.4.4. Description des articles et des fichiers logiques .....	148.
2.4.5. Organisation physique des données .....	150.
2.4.5.a. Format interne .....	150.
2.4.5.b. Format externe d'entrée .....	152.
2.4.5.c. Format externe de sortie .....	152.
2.4.6. Organisation physique des fichiers .....	152.
2.4.6.a. Première possibilité : organisation type-1 .....	153.
2.4.6.b. Deuxième possibilité : organisation séquentielle .....	157.
2.4.6.c. Troisième possibilité : organisation séquentielle avec fichiers inversés .....	159.
2.4.7. Evaluation des performances globales .....	160.
2.4.7.a. Les modélisations .....	160.
2.4.7.b. Les quantifications .....	160.
2.4.7.c. Caractéristiques physiques des mémoires de masse .....	164.
2.4.7.d. Projet .....	165.
2.4.7.e. Hypothèses .....	166.
2.4.7.f. Estimation des temps .....	166.
2.4.7.g. Premier commentaire .....	168.
2.4.7.h. Prise en compte du temps de réponse global (accès disque) .....	168.
h.1. Un seul élément demandé .....	169.
h.2. Plusieurs éléments demandés .....	169.
2.4.7.i. Choix d'une organisation définitive. Conclusion .....	171.
2.5. Analyse des traitements .....	171.

2.5.1. Enregistrement .....	171.
2.5.2. Mise à jour .....	171.
2.5.3. Suppression .....	172.
2.5.4. Interrogation. Consultation .....	172.
2.5.4.a. Consultation d'un dossier sur base de son numéro .....	173.
2.5.4.b. Requêtes spécifiques sur base de cri- tères définis par l'utilisateur .....	174.
2.5.4.c. Etablissement de séries de fréquences ..	176.
2.5.5. Production d'états statistiques .....	177.
§3. Le programme d'aide au diagnostic .....	177.
3.1. L'ancien système .....	178.
3.2. Le système d'AIDE au DIAGNOSTIC proposé (AIDDIA) .....	178.
3.2.1. Routine TRIMOD .....	180.
3.2.2. Routine DIAMOD .....	180.
3.2.3. Routine UTIMOD .....	180.
3.2.4. Routine OPTMOD .....	181.
3.2.5. Routine LISTMOD .....	181.
3.2.6. Routine CTSMOD .....	182.
3.3. Remarques .....	182.
 Conclusion générale .....	 185.
 Bibliographie .....	 187.
 Annexes .....	 195.



## INTRODUCTION GENERALE.

Ce mémoire poursuit essentiellement deux objectifs.

Le premier consiste à établir un bilan des applications de l'informatique à la médecine, et plus spécialement au diagnostic médical.

Nous nous efforcerons d'y parvenir en détaillant certains modèles théoriques puis en les illustrant, en mettant également en lumière un certain nombre de problèmes y attachés.

Le second concerne une situation concrète. Il consiste en la conception et l'analyse détaillée d'un système d'information destiné au centre anti-poisons de Bruxelles.

La réalisation de ce deuxième objectif nous verra partir d'une critique de l'existant et d'une mise en lumière des besoins justifiant l'implantation de ce système, pour aboutir à la description détaillée des spécifications du système que nous concevrons.

Ces deux objectifs justifient par ailleurs la découpe de notre étude en deux parties.

Ière PARTIE : L'AIDE AU DIAGNOSTIC MEDICAL.

METHODES ET APPLICATIONS.



## INTRODUCTION.

Dans ce premier volet de notre étude, nous prendrons conscience du problème qui nous est posé, à savoir l'introduction de l'outil informatique dans les milieux médicaux. Nous passerons rapidement en revue les différents domaines d'application actuellement explorés avant d'étudier plus en détail les difficultés inhérentes à la cohabitation médecine-informatique. Nous enchaînerons par la mise en exergue de trois problèmes de fond qui restent à l'heure actuelle au coeur d'un débat qui met aux prises praticiens et informaticiens.

Au-delà de quelques notions terminologiques, nous résumerons alors en quelques mots la problématique du diagnostic pour tenter de la formaliser. Au plan plus particulier de l'aide au diagnostic, nous analyserons les avantages offerts par l'automatisation ainsi que les limites qu'il faut lui fixer.

Nous entrerons ensuite dans une partie plus théorique qui nous verra présenter un éventail assez large des méthodes d'aide au diagnostic les plus couramment employées. Ce chapitre de méthodologie sera complété par l'exposé de différentes stratégies de choix d'examen complémentaires, techniques souvent combinées avec les précédentes dans les applications d'aide au diagnostic.

Nous illustrerons alors l'exposé des méthodes à l'aide de deux exemples d'applications sur lesquelles nous avons eu l'occasion de nous attarder durant notre stage à Paris, deux exemples qui visent essentiellement à montrer deux études originales basées, l'une sur l'analyse de données multidimensionnelle, la seconde sur la méthode probabiliste. L'exposé qui en sera fait outrepassera l'application proprement dite d'un algorithme pour détailler l'ensemble de chaque projet.

Enfin, nous nous attacherons à donner un avis plus personnel sur les perspectives d'avenir que nous avons pu discerner tout au long de notre travail, de nos enquêtes, en faisant le point des réalisations, des échecs et des espoirs entretenus.



## CHAPITRE 1. INFORMATIQUE ET MEDECINE.

### A. Prise de conscience du problème : une ère d'évolution.

Depuis plus d'un demi siècle, le développement des techniques de recherche a profondément accéléré le processus d'évolution de tous les domaines scientifiques. Le domaine médical n'a pas échappé à la règle. Il est, lui aussi, entré de plein pied dans cette ère de bouleversement constant, où chaque jour apporte sa contribution à la "Science", à la "Connaissance de l'homme". L'accumulation des études, des découvertes a rendu nécessaire l'utilisation d'un outil de maîtrise, de diffusion et d'exploitation de cette information nouvelle.

D'un autre côté, le perfectionnement des ordinateurs entraînant la mise à disposition immédiate de volumes incroyables de données complexes et le développement des systèmes de gestion des bases de données ont offert un outil particulièrement intéressant.

Le pas fut donc bien vite franchi, maintes applications imaginées, certaines concrétisées pour constituer ce qu'aujourd'hui nous appelons, d'un terme fort ambitieux, l'informatique médicale.

### B. Les domaines actuels d'application.

Les réalisations actuelles, dans le cadre de l'informatique médicale, s'étendent à différents domaines, parmi lesquels les plus courants sont devenus :

- la gestion du dossier médical,
- le dépouillement des revues et travaux scientifiques,
- la gestion administrative de l'hôpital,
- la surveillance des malades par monitoring,
- l'analyse des examens paracliniques,
- la transmission des résultats de laboratoire,
- l'aide au diagnostic.

Il semble clair que la plupart de ces applications relèvent davantage du domaine de la gestion que de la médecine en tant que pratique médicale, qui a provoqué bien plus de réticences, de son côté.

### C. Les difficultés de mise en place.

En effet, comme toute révolution, l'introduction d'un système informatique dans le monde médical a provoqué des levées de résistance, tantôt aveugles, tantôt justifiées. Les principaux motifs avancés se résument aux suivants :

- le succès douteux et partiel des premières tentatives,
- l'impression qui s'est dégagée d'une complexification des problèmes abordés qui ne concerneraient plus que des experts, à l'exclusion des médecins non-initiés,
- la nécessité d'une certaine restructuration des services où l'ordinateur serait implanté et l'obligation d'y resté lié et soumis,
- un coût d'acquisition et de mise en place important sans garantie ou preuve des services et avantages à en dégager (il est difficile de parler de profit),
- la perspective d'une période assez longue de mise en place et d'adaptation au système,
- la crainte des problèmes que ne manquerait de soulever une défaillance du système,
- la façon de penser de certains médecins qui y présentent un moyen de contrôle de leur activité ainsi que la perte d'un certain prestige séculier.

Pour répondre dès à présent à certaines de ces objections et dans l'espoir de faire accepter cette intrusion, il nous semble indispensable de :

- préparer psychologiquement à une compréhension des avantages apportés et établir un changement progressif des habitudes de travail,
- souligner la libération de nombreuses tâches fatigantes,
- mettre en place rapidement les fonctions essentielles pour limiter les hésitations, les remises en cause



- conserver le contrôle de tout ce qui est entré en machine pour ne pas lui être soumis mais l'utiliser comme un outil de travail,
- prévoir les incidents possibles pour limiter les perturbations et garantir la sécurité et la disponibilité du système.

D. Problèmes de fond : le secret médical, le conflit de compétence et la responsabilité.

Au-delà de ces différents aspects, non négligeables, subsistent trois autres sujets de polémiques, plus fondamentaux, très souvent avancés de manière plus catégorique encore par les opposants à l'informatisation de la médecine. Nous voulons parler du secret médical, du conflit de compétences et du problème de la responsabilité, ces deux derniers points s'adressant plus spécifiquement aux différents systèmes d'aide au diagnostic. Ce sont trois problèmes réels, trois pierres d'achoppement sur lesquelles il convient d'adopter de prime abord une position non ambiguë qui soit la conclusion d'un débat entre les médecins, les informaticiens et le législateur.

§1. Le secret médical.

A l'heure actuelle, le secret médical, défini par la loi, bien qu'attaqué et mal protégé, ne peut être violé que dans certains cas limités. Dans les hôpitaux, la protection du secret des informations contenues dans les archives médicales, n'est pas organisée systématiquement. Ni la qualification du personnel, ni l'infrastructure matérielle ne le permettent dans la plupart des cas. En quoi le traitement des dossiers médicaux par ordinateur qui s'étend lentement mais, semble-t-il, implacablement, peut-il donc modifier la situation ?

Il est évident que certaines des caractéristiques du traitement informatique sont favorables à une meilleure protection du secret des données médicales.

Nous pensons à toutes les techniques dites de sécurité, constamment remises en cause et repensées. Mais citons à ce propos l'avis émis par J. Berleur et Y. Pouillet (B8, B9).

"L'intérêt de ces techniques est de mettre aux mains des gestionnaires les moyens de répondre positivement à l'intention et aux dispositions de ceux qui réclament une protection maximale contre toute lecture ou traitement abusif de données individuelles. Mais, -nous voudrions insister sur ce point-, elles ne créent que les conditions de possibilité d'une vraie protection et ne peuvent, en aucune manière, être l'alibi d'un attentisme satisfait ou d'un laisser-faire de mauvais aloi : il serait illusoire d'espérer que le choix puisse se faire sur une base purement technique."

Si le problème se pose de manière si cruciale, c'est parce que l'organisation structurée, la mise à disposition, la centralisation des données présentent tout autant un risque de diffusion non contrôlée réel qu'un intérêt évident pour le malade comme pour le médecin. Combien de fois un médecin n'a-t-il pas déploré de ne pouvoir obtenir sans délais les renseignements médicaux relatifs au patient inconnu qui lui était présenté ? Que de doutes entretenus, que de temps gaspillé !

Face à ce problème, nous prôtons, nous aussi, à l'image de beaucoup, l'extension du secret médical à toute personne qui s'engage dans le milieu médical, de quelque manière que ce soit. Le secret s'étendrait donc aux informaticiens concernés, au personnel de saisie comme au personnel des services de soins. J. Berleur et Y. Pouillet affirment aussi, traitant plus généralement de la protection de la vie privée :

"Les gestionnaires des Centres d'informatique peuvent être tenus pour pleinement responsables de la sécurité de leurs données. Obligation pourrait leur être faite



d'assumer cette responsabilité, au moins quant aux moyens, mais aussi quant aux résultats..."

## §2. La compétence du médecin.

Sans aucun doute, le diagnostic automatique pose un problème de conflit de compétences. S'il nous est permis de penser qu'un système en place nivellerait les valeurs par le haut, en mettant à la portée des non-spécialistes et des débutants la connaissance des plus expérimentés, nous pouvons aussi pressentir que le recours de l'intelligence humaine à ces méthodes d'assistance contribuerait à ternir l'image du médecin vis-à-vis de ses malades.

Mais nous sommes loin du compte, loin du jour, pensons-nous, où la machine posera et imposera son diagnostic, si toutefois il doit venir.

## §3. La responsabilité.

Qui va, dans l'optique d'un système d'aide au diagnostic automatique ou dans tout autre système à vocation médicale, supporter la responsabilité des défaillances matérielles ou des erreurs de saisie ? Est-il concevable d'envisager un partage de cette responsabilité entre le médecin, les concepteurs et les gestionnaires du système ?

Ces deux questions débouchent, elles aussi, sur un débat au terme duquel deux tendances se font jour. L'une prône une responsabilité collégiale ; mais si tous sont responsables, personne ne se sent directement mis en cause. Nous estimons que cette tendance vise à rejeter les erreurs sur un ensemble, sur un système qui finira par être accablé. La seconde, et c'est là notre opinion, rejoint l'idée de M. Lortat et M. Jacob (réf.) qui affirmaient :

"Le progrès des techniques qui a décuplé les possibilités de la médecine ne change pas la nature de la

responsabilité du médecin."

Toujours est-il, pour développer notre pensée, que le diagnostic automatisé ne doit être considéré que comme une assistance au diagnostic médical que le médecin est seul à pouvoir porter sous son entière et unique responsabilité. L'outil doit essayer d'approcher la réalité du patient avec le plus de certitude possible et présenter au médecin des hypothèses sujettes à son interprétation. Le nombre de facteurs et les circonstances extrinsèques sont tellement nombreux et imprécisables qu'il s'avère utopique d'oser croire en l'avènement d'une médecine sans médecins.

---

LORTAT - JACOB J.L. -VILLEY R. et GUENIOT M. (1968)

'Responsabilité personnelle du médecin.' In 2<sup>e</sup> Congr. int. morale médicale, Paris 1966, Ordre National des Médecins éd.



## CHAPITRE 2. L'AIDE AU DIAGNOSTIC MEDICAL.

### A. Caractéristiques générales.

#### §1. Quelques définitions.

Un symptôme : caractéristique observable chez un malade.

Un signe : caractéristique observée chez un malade. C'est la valeur prise par un symptôme.

Une symptomatologie : ensemble de symptômes.

Un syndrome : ensemble de signes.

Un diagnostic : caractéristique inobservable d'un malade. De ce fait, ce que l'on appelle couramment diagnostic correspond à des niveaux de connaissance très divers. Un médecin pense avoir posé un diagnostic lorsqu'il a, selon les circonstances :

- reconnu simplement l'existence du symptôme,
- reconnu son appartenance à un syndrome,
- découvert une maladie de cause inconnue,
- découvert une maladie de cause connue.

Dans le sens que nous considérons ici, le diagnostic est pratiquement équivalent à la maladie.

Une nosographie : ensemble de diagnostics (partition).

#### §2. La problématique du diagnostic médical.

La problématique du médecin face à son patient se résume en une alternative :

- soit il estime que les renseignements dont il dispose suffisent pour l'amener à poser son diagnostic avec un niveau de certitude acceptable,
- soit, jugeant ces renseignements insuffisants, il prescrit des examens complémentaires susceptibles, selon son opinion, de confirmer ou d'infirmer son diagnostic. Le choix de ces examens doit alors répon-

dre à différents critères tels que :

- minimiser le temps d'attente,
- minimiser le nombre de tests nécessaires pour aboutir à la formulation d'un diagnostic satisfaisant,
- minimiser le coût global de ces examens,
- minimiser le désagrément et les risques encourus par le malade, compte tenu de son état de santé actuel, de sa tolérance envers certains examens, de son état psychologique, ...etc.

### §3. Schématisation de la démarche diagnostique.

Nous pouvons dissocier deux démarches différentes.

#### 3.1. Le diagnostic global.

Il s'apparente à la "reconnaissance des formes", c'est-à-dire à la prise en compte globale, instantanée de l'état du patient, qui pousse instinctivement le médecin à établir son diagnostic, sans s'attarder à l'identification d'un syndrome.

#### 3.2. Le diagnostic pas à pas.

Chaque étape du raisonnement est constituée d'une question dont la ou les réponses conditionnent le choix de la question suivante. Ce processus pas à pas ne se présente pas comme un processus linéaire strict, dans la mesure où les signes sont recueillis par groupes, leur contribution et leur prise en considération restant totalement aléatoires. L'orientation vers la question suivante est fonction de l'interprétation que fait chaque clinicien du syndrome présenté, associé à sa modalité de prise en compte, totalement instable et subjective, des critères énoncés au point précédent. Sans oser prétendre à la standardisation d'une démarche diagnostique, il nous faut cependant tenter de schématiser



ce processus, en faisant provisoirement abstraction, ce qui ne revient pas à désavouer leur pertinence, de cette foule d'éléments subjectifs.

#### Formalisation de la démarche pas à pas.

- 1- Interrogation du malade qui débouche sur la reconnaissance d'un syndrome.
- 2- Orientation vers l'un ou l'autre groupement de diagnostics.
- 3- Recherche de signes cliniques ou paracliniques présents dans les affections supposées à ce stade.
- 4- Examen minutieux et confrontation des signes aux formes-types des différents diagnostics envisagés.
- 5- Soit le diagnostic est établi,  
soit recherche des formes cliniques (où manque un signe, puis deux signes,...) des diagnostics supposés. Cette étape inclut principalement les problèmes délicats posés par l'existence de formes cliniques presque identiques à des affections différentes (notion de diagnostic différentiel). Retour au point 4.

### B. Automatisation de l'aide au diagnostic.

#### §1. L'objectif.

L'objectif primaire se définit clairement : accroître le pourcentage de diagnostics exacts tout en augmentant leur degré de certitude.

#### §2. L'apport de l'informatique.

L'aide de l'informatique peut intervenir à diverses étapes :

- lors de l'acquisition de données qui doivent conduire à l'établissement du diagnostic, en vue de l'optimisation du choix des examens complémentaires,
- lors de l'interprétation de ces données, pour permettre la confrontation des hypothèses du médecin et des résultats (probabilités,...) présentés par le calculateur. Une diver-

gence de vue peut amener progressivement le médecin à mieux apprécier la valeur des informations précédemment recueillies (ou à critiquer leur validité). On peut arriver ainsi à une remise en question permanente (mais mesurée) des données stockées et à un auto-contrôle du clinicien. Le contrôle implacable opéré par la machine influence la pratique quotidienne, éliminant les imprécisions et les doutes dûs à la fatigue, au caractère habituel et quasi-routinier de certaines prises de décisions.

D'autre part, l'ordinateur permet de synthétiser l'expérience des meilleurs experts pour la mettre à la portée de tous. Enfin, les systèmes informatisés d'aide au diagnostic peuvent constituer un outil précieux en vue de l'enseignement. Cependant, il nous paraît bon de placer ici un avertissement fait à ce propos, qui n'est peut-être pas partagé par tous :

"L'aide à la décision médicale doit s'individualiser

- des systèmes destinés à l'auto-apprentissage,
- de la recherche,

sous peine de devenir des modèles inadaptés, donc inefficaces." H.Ducrot (B17).

Cette opinion ne dénie pas l'utilisation possible d'un outil d'aide au diagnostic à d'autres fins, mais refuse le fait que, lors de la conception d'une application précise, des intérêts secondaires entravent son développement et nuisent à l'efficacité de l'objectif majeur.

### §3. Les limites de compétence.

Bien que les diverses perspectives énumérées laissent entrevoir la concrétisation de nombreux espoirs, il faut dès à présent restreindre le champ des applications. Pour disposer d'un cadre plus favorable à une étude moderne des méthodes de diagnostic et à une tentative de standardisation des thérapeutiques, il faut :

- que le processus de diagnostic réponde à une logique modélisable qui n'intègre pas d'éléments subjectifs,



- un nombre d'affections possibles relativement limité (ou une réduction du nombre des affections à un nombre restreint de groupes d'affections),
- une symptomatologie aisément quantifiable et standardisable,
- que les procédés diagnostiques employés amènent des résultats chiffrés, aisément manipulables par les techniques statistiques.

D'un autre côté, la symptomatologie, toute évocatrice qu'elle soit, n'est pas forcément décisive. Il existe des variations de symptomatologie selon l'âge, le sexe, le stade de la maladie et bien souvent, on ne dispose pas d'un échantillon suffisant de chaque cas possible pour en tirer des chiffres fiables.

## CHAPITRE 3. METHODOLOGIE.

### A. Les méthodes d'aide au diagnostic.

#### Avertissement.

Dans cette partie essentiellement théorique de notre étude, nous passerons en revue les principes inhérents aux méthodes actuellement appliquées dans certaines applications d'aide au diagnostic. Ce recueil ne se veut pas exhaustif mais il tente de présenter, de façon plus ou moins objective, ce qui a été découvert au cours des stages. Aussi avons-nous décidé d'uniformiser la présentation de ces méthodes. Le texte en marge duquel figure un trait est donc le fruit d'un travail commun à Th. Biernaux et P. Deroy. (réf.)

#### §1. Le diagnostic global.

##### 1.1. Les méthodes d'analyse multidimensionnelle.

Ces techniques présentent un certain intérêt pour :

- la mise en évidence de certains aspects de l'association entre variables,
- l'étude des relations entre un paramètre (ou plusieurs) et un ensemble d'autres paramètres,
- la partition d'un ensemble d'observations en groupes homogènes,
- la détermination du 'si' et du 'comment' deux ou plusieurs groupes diffèrent dans leur profil multidimensionnel.

La notion fondamentale de toutes ces techniques est celle

---

DEROY P. (1979)

'Aide au diagnostic médical. Synthèse des méthodes et applications.' , mémoire, Institut d'Informatique-Namur.



de distance servant à définir la proximité ou la ressemblance existant entre observations ou paramètres.

Cette notion de distance distingue différentes méthodes d'analyse multidimensionnelle, parmi lesquelles il nous faut citer :

- l'Analyse des Correspondances,
- l'Analyse en Composantes Principales,
- l'Analyse Discriminante Linéaire ou Quadratique.

Comme le développement théorique de ces principales méthodes applicables ne présente que peu d'intérêt dans le cadre de notre étude, nous renvoyons le lecteur plus intéressé à trois ouvrages :

- BENZECRI J.P. 'Analyse des Correspondances'  
tome II DUNOD 1973.
- DRAPER N.R., SMITH H. 'Applied Regression Analysis'  
WILEY 1967.
- SEBESTYEN G.S. 'Decision Making Process in Pattern Recognition' THE MAC MILLAN COMPANY 1962.

\* principe :

Ces méthodes procèdent généralement en deux étapes. La première utilise un échantillon de malades de diagnostic connu dont elle analyse les observations et parfois même les paramètres (Analyse des Correspondances). La seconde étape prend en compte un nouveau cas clinique qu'elle confronte à la connaissance scientifique tirée de l'échantillon et synthétisée, en vue d'en établir le diagnostic ou le pronostic.

#### 1.1.a. Méthodes descriptives.

La méthodologie de l'analyse factorielle est basée sur la recherche d'un système réduit de références généré par de nouveaux axes. Ces nouveaux axes, appelés axes factoriels, sont obtenus par une rotation dans l'espace à  $p$  dimensions du système de coordonnées initial dans une orientation cor-



respondant aux directions de variance maximum. Les axes factoriels sont choisis en ordre décroissant.

Il suffit alors de ne considérer que les axes factoriels représentatifs d'un pourcentage suffisant de la variance totale, ce qui entraîne une réduction de la dimension du nouveau système.

Plusieurs méthodes relèvent de la même démarche :

- l'Analyse en Composantes Principales,
- l'Analyse des Correspondances.

#### 1.1.b. Méthodes de classification.

Si l'analyse factorielle a révélé l'existence de groupes et si le pourcentage de variance totale prise en compte par les trois premiers axes n'est pas significativement important, il faut approfondir l'analyse de ces données et essayer de confirmer ces résultats par des méthodes de classification ou de groupages.

Leur but est de définir une partition des éléments de l'échantillon en groupes homogènes et distincts.

On peut distinguer les méthodes hiérarchiques des méthodes non hiérarchiques.

##### - classification hiérarchique :

Cette analyse permet de construire un arbre de classification (dendogramme) dont les différents niveaux représentent une partition de l'échantillon. Cet arbre sert à classer les patients et les paramètres et procède en trois étapes :

- 1- Codification des données.
- 2- Choix de la distance adéquate entre les éléments à classer.
- 3- Choix d'un algorithme d'agrégation.

##### - classification non hiérarchique ou dynamique :

Elle débouche sur une partition optimale des éléments à classer. Le nombre  $k$  de groupes est fixé à l'avance, car c'est la base de la procédure, certaines classes pouvant être vides si  $k$  a été surestimé.

\* procédure :

- k éléments tirés au hasard de l'échantillon et appelés 'éléments standards' constituent les noyaux des k classes.
- On calcule les distances entre chaque élément de l'échantillon et les différents éléments standards, puis on alloue chaque élément au noyau le plus proche. On obtient donc k classes disjointes.
- Dans chaque classe, on prend un élément au hasard, qui constitue un nouvel élément standard. Un nouveau calcul des distances fournit une génération de k nouvelles classes disjointes.
- On itère ainsi le processus jusqu'à stabilité de la partition.

#### 1.1.c. Méthodes décisionnelles.

L'un des paramètres joue ici un rôle spécial; c'est la variable dépendante, que l'on oppose aux autres variables qui constituent l'ensemble des variables indépendantes ou explicatives.

\* méthode :

- On détermine, d'après l'échantillon, si la variable dépendante est significativement expliquée par les autres.
- Si la première étape révèle une relation significative on peut prédire la valeur de la variable dépendante pour une nouvelle observation.

Parmi les principales méthodes décisionnelles applicables, selon la nature des variables (continue, discrète, binaire), citons : - l'Analyse des Correspondances,  
 - l'Analyse en Composantes Principales,  
 - l'Analyse Discriminante Linéaire,  
 - l'Analyse Discriminante Quadratique.



### 1.1.d. Méthodologie de l'analyse des correspondances.

Pour mettre en lumière les principes énoncés, et en pré-lude à l'exemple d'application que nous présentons au chapitre quatrième de notre étude, nous allons exposer la méthodologie de l'analyse des correspondances, développée par Fisher puis surtout par Benzécri.

C'est une méthode à la fois descriptive et explicative qui débouche sur des règles d'interprétations très simples.

Considérons  $T$ , la table des données à analyser,

$J$ , le vecteur des paramètres (les colonnes de  $T$ ),

$I$ , le vecteur des observations (lignes de  $T$ ).

Les données  $t(i,j)$  sont d'abord traduites en fréquences  $y(i,j)$  de manière à rendre les observations comparables.

$$y(i,j) = \frac{f_{ij}}{f_{i.}} \quad i \in I \text{ et } j \in J$$

$$\text{où } f_{ij} = \frac{t(i,j)}{\sum_i \sum_j t(i,j)}$$

$$f_{i.} = \sum_j f_{ij}$$

On procède ensuite au calcul de la distance CHI-CARRE entre observations :

$$d^2(i,k) = \sum_j \frac{1}{f_{.j}} \left[ \frac{f_{ij}}{f_{i.}} - \frac{f_{kj}}{f_{k.}} \right]^2$$

La recherche du nouveau système de référence généré par les axes factoriels correspondant aux directions de variance maximum passe par la résolution de l'équation :

$$M V \varphi = \lambda \varphi$$



où  $M$  est la matrice des distances calculées,

$V$  est la matrice de variance-covariance calculée à partir des  $y(i,j)$ ,

$\phi$  et  $\lambda$  sont respectivement les vecteurs propres orthonormés et les valeurs propres associées de la matrice  $MV$ .

Les coordonnées des observations sur l'axe factoriel  $F_r$  s'obtiennent par :

$$\psi_{ir} = \frac{1}{\sqrt{\lambda_r}} * \sum_j \frac{f_{ij}}{f_{i.}} * \phi_{jr}$$

où  $\phi_{jr}$  représente le  $j$ -ème composant du  $r$ -ème vecteur propre.

En appliquant la même démarche au vecteur des paramètres, on obtient les formules de transition suivantes :

$$\begin{aligned} \psi_{ir} &= \frac{1}{\sqrt{\lambda_r}} * \sum_j \frac{f_{ij}}{f_{i.}} * \phi_{jr} \\ \psi_{jr} &= \frac{1}{\sqrt{\lambda_r}} * \sum_i \frac{f_{ij}}{f_{.j}} * \phi_{ir} \end{aligned}$$

Ces formules de transition nous permettent donc de représenter les observations et les paramètres sur un seul et même plan. La proximité d'une observation et d'un paramètre sur ce plan traduit le fait que cette observation est particulièrement bien expliquée par ce paramètre. Cela nous permet d'établir des groupes.

Toute nouvelle observation (ou nouveau paramètre) peut alors être reportée sur le plan grâce à l'application des formules de transition, et sa position est interprétée plus ou moins aisément, selon les cas.

\* conclusion :

En résumé, selon la nature des données utilisées, il s'agit d'employer :

- la méthode d'analyse factorielle adéquate en vue d'obtenir une description générale des données et, si possible, une réduction de ces données,
- la méthode de classification adéquate pour confirmer un éventuel groupage de patients révélé par l'analyse factorielle,
- la méthode explicative (ou décisionnelle) adéquate, pour discriminer les groupes en obtenant une règle de prise de décision.

\* critique :

Les méthodes d'analyse de données multidimensionnelle n'utilisent que des variables quantitatives (résultats de dosage,...). Elles ne permettent pas de prendre directement en compte les signes cliniques, discrets et qualitatifs, qui se bornent à déterminer la présence ou l'absence d'un attribut.

D'un autre côté, elles ont la manie de classer tous les cas envisagés, si bien qu'une affection évidente se trouve subitement entachée d'un doute.

## 1.2. Les méthodes de discrimination non paramétrique.

\* principe :

A partir d'échantillons de malades de diagnostic connu, on se propose d'établir des règles de classement sans référence à des lois de probabilité.

Ce type de méthode est utilisé dans des cas très particuliers.



\* exemple : Méthode de diagnostic par index clinique.

On recherche un certain nombre de signes considérés comme étant de poids diagnostique important. Ces signes sont affectés d'une pondération plus ou moins arbitraire.

On additionne les poids des signes dont on relève la présence pour établir la valeur d'un index.

Une étude préliminaire sur un fichier de malades cliniquement diagnostiqués permet de connaître la répartition des valeurs de l'index ainsi établi en fonction des diagnostics.

Il existe parfois des zones de recouvrement où l'établissement d'un diagnostic reste douteux.

Cette méthode a été appliquée par de nombreux thyroïdologues. Il s'avère que l'index le plus simple, mis au point par R. Simonin (B56) en 1966 (pondération des signes = 1 si le signe est présent et = 0 sinon), est souvent le plus performant.

\* critique :

Cette méthode ne peut s'appliquer que dans certains cas très spécifiques, pour des pathologies où le nombre de symptômes et de diagnostics est limité.

De plus, elle est mise en défaut par la présence d'un symptôme rare.

## §2. Le diagnostic pas à pas.

### 2.1. Les méthodes logiques.

Ces méthodes se basent essentiellement sur la logique booléenne. Mais elles dépassent ce cadre jugé trop contraignant en s'étendant, comme nous allons successivement le montrer, à la logique ternaire et aux principes régissant la démonstration automatique des théorèmes.

#### 2.1.1. Application du calcul booléen.

Cette optique tente de classer le malade, d'abord dans un cadre nosologique assez large, puis dans des cadres de plus



en plus restreints, au fil des pas diagnostiques.

Lorsque le médecin consulte le malade, il ne peut prendre en compte simultanément plus de quelques signes. La méthode logique adopte cette démarche en utilisant, à chaque itération, un nombre restreint de signes différents.

Pour mettre en lumière ces pas diagnostiques, nous allons distinguer deux phases, la première explicative, la seconde de classification.

- phase explicative :

Il s'agit de la phase la plus complexe de la démarche. En effet, il faut sélectionner, parmi tous les signes susceptibles d'être recueillis, ceux qui orientent le mieux vers tel ou tel diagnostic. Lorsque ces signes et/ou examens ont été choisis, il faut trouver leurs associations qui déterminent des groupes diagnostiques les plus précis possible.

Deux méthodes sont utilisées pour la phase explicative :

- la première est basée sur le test de CHI-CARRE.

(CHU Timone de Marseille)

- la seconde s'inspire de la théorie des graphes.

(Begon, Trémolières)

- phase de classification :

Cette phase de classification est assez simple car elle se borne à l'application des fonctions logiques obtenues aux signes et/ou examens des nouveaux patients à diagnostiquer.

2.1.1.A. Phase explicative.

a. Méthode basée sur le test de CHI-CARRE.

a.1. Sélection des signes.

Soit S, un ensemble de signes cliniques systématiquement recueillis par un système de questionnaire. De cet ensemble, il nous faut extraire ceux dont le poids diagnostique est le plus fort. Cette sélection s'obtient par le calcul du CHI-CARRE opéré pour chaque signe sur une table de con-

tingence du type suivant :

	$D_k$	$\bar{D}_k$	
$S_1$	$n_1$	$n_3$	$n_1 + n_3$
$\bar{S}_1$	$n_2$	$n_4$	$n_2 + n_4$
	$n_1 + n_2$	$n_3 + n_4$	$n_1 + n_2 + n_3 + n_4$

où  $D_k$  représente le diagnostic k,  
 $\bar{D}_k$  représente l'absence du diagnostic k,  
 $S_1$  représente la présence du signe 1,  
 $\bar{S}_1$  représente l'absence du signe 1,  
 $n_i$  représente le nombre de patient de diagnostic k ou de diagnostic différent de k, qui présentent ou qui ne présentent pas le symptôme 1.

La valeur du CHI-CARRE détermine ainsi le poids diagnostique de chaque signe.

Cependant, il peut arriver que deux signes sélectionnés de cette façon apportent la même information pour un diagnostic donné. Il faut alors, par un examen complémentaire, éliminer, parmi les signes qui offrent le plus grand lien avec chaque diagnostic, ceux qui s'avèrent être fortement liés entre eux.

Cela peut se faire par le calcul de la distance de Hamming (poids du vecteur obtenu par l'application de l'opérateur 'ou exclusif' aux vecteurs d'observations relatifs à deux patients, ce qui détermine en fait le nombre de signes différents de l'un à l'autre).

On divise alors l'échantillon de départ en deux sous-ensembles, suite aux réponses positives ou négatives faites à certains signes. Parmi ces signes, dits signes de séparation, on recherche ceux qui minimisent la distance intra-groupe, c'est-à-dire la somme des distances entre un patient et les autres du même groupe, et qui maximisent la distance inter-groupes, c'est-à-dire la distance entre un patient et tous les patients de l'autre groupe.



### a.2. Sélection des associations.

Il faut ici insister sur le fait que, dans ces associations, tous les signes importent. Il ne suffit donc pas de sélectionner uniquement les signes positifs les plus fréquents. Les signes négatifs ont un poids identique.

Sur base des dossiers, il faut rechercher les implications du type suivant :

$m_i \rightarrow D_k$  où  $D_k$  est le k-ième groupe diagnostique au pas considéré,  
 $m_i$  est le minterm construit sur un ensemble de signes.

Il faut ainsi trouver les minterms qui ne correspondent qu'à un seul groupe diagnostique.

Pour les découvrir, on construit un tableau associé à chacun des diagnostics possibles. On explore le dossier médical en remplissant les différentes valeurs, positives ou négatives, associées aux différents signes retenus lors de la première étape du processus, à savoir la recherche des signes les plus discriminants.

Par diagnostic, on obtient un tableau du type suivant :

$D_1$	$S_1$	$S_2$	$S_3$	$S_4$	..	$S_i$	..
patient 1	1	0	0	1	..	1	..
patient 2	1	0	1	0	..	1	..
patient 3	1	1	0	1	..	1	..
...							
patient m	..	..	..	..	..	..	..

où une colonne représente les valeurs d'un signe chez les différents patients,  
une ligne représente les symptômes présentés par un patient.

Pour trouver les minterms spécifiques à un diagnostic, on élimine les associations de signes, c'est-à-dire les combinaisons de 0 et de 1 qui sont communes à ces différents



tableaux.

C'est ici que l'on se rend compte de l'utilité de l'hypothèse d'exclusivité des diagnostics; elle permet en effet d'éviter la constitution des tableaux relatifs aux combinaisons de diagnostics.

Les combinaisons qui subsistent alors dans le tableau relatif à chaque diagnostic sont donc spécifiques à ce diagnostic.

Pour les  $k$  groupes diagnostiques du pas considéré, on a ainsi sélectionné un certain nombre de minterms qui vérifient l'implication  $m_i \rightarrow D_k$ . On peut alors construire, par groupe diagnostique  $k$ , une fonction  $F_k$ :

$$F_k = \sum_i m_i \quad i \in I, \text{ ensemble des minterms qui vérifient la première relation.}$$

Cette fonction adopte une forme disjonctive.

L'ensemble des associations de signes identifiées peuvent se représenter sous forme d'un tableau :

SIGNES DISCRIMINANTS	$S_1$				
	$S_2$	$m_1$	$m_2$	..	$m_i$
	$S_3$				
	..				
	$S_m$				
DIAGNOSTICS POSSIBLES	$D_1$	X	X		
	$D_2$			X	
	..				
	$D_n$				X

# b. Méthode basée sur la théorie des graphes.

Comme pour la méthode utilisée au CHU Timone de Marseille, elle a pour objectif de proposer des signes et des associations de signes et/ou d'examens impliquant des groupes de diagnostics.

Nous avons déjà soulevé au point précédent, le problème du choix des signes. Dans cette optique, nous traiterons du choix sous critère. En effet, chaque examen s'accommode d'inconvénients dont il nous faut tenir compte, inconvénients de coût, de durée, ...etc.

Nous posons la même hypothèse d'exclusivité des diagnostics.

Nous développerons un exemple théorique qui met en jeu cinq diagnostics, six examens et deux critères de coût.

Soient  $E_j$ , les examens considérés,

Soient  $D_i$ , les cinq diagnostics possibles.

CRITERES		TABLE de DIAGNOSTICS					
COUT (en B)	TEMPS (en minutes)	EXAMENS	MALADIES				
40	4	$E_1$	0	0	0	0	1
15	7	$E_2$	1	0	0	0	0
50	4	$E_3$	1	1	1	0	1
16	5	$E_4$	0	0	0	1	1
75	20	$E_5$	0	1	0	0	1
8	5	$E_6$	1	0	1	0	0
			$D_1$	$D_2$	$D_3$	$D_4$	$D_5$

Sur ce tableau, une ligne montre le profil d'un examen pour les différents diagnostics tandis qu'une colonne traduit le profil d'un diagnostic au travers des différents examens.



## \* principe :

La démarche diagnostique proposée par Begon-Trémolières consiste à déterminer tous les sous-ensembles d'examens permettant de distinguer les différents diagnostics (ces ensembles étant appelés ensembles filtrants), et à choisir, en fonction de certains critères, l'un de ces sous-ensembles.

## \* méthode :

Dans un premier temps, on va chercher tous les sous-ensembles d'examens qui différencient à la fois  $D_1$  de  $D_2$  et  $D_1$  de  $D_3$  et ... et  $D_4$  de  $D_5$ , ce qui donne le tableau suivant, appelé tableau des ET :

POUR DISTINGUER	IL FAUT FAIRE LES EXAMENS	
ET ( $D^1$ de $D^2$ )	$E_2$ ou $E_5$ ou $E_6$ ) ET	$L_1$
ET ( - de $D^3$ )	$E_2$ ) ET	$L_2$
ET ( - de $D^4$ )	$E_2$ ou $E_3$ ou $E_4$ ou $E_6$ ) ET	$L_3$
ET ( - de $D^5$ )	$E_1$ ou $E_2$ ou $E_4$ ou $E_5$ ou $E_6$ )	$L_4$
. $D^2$ de $D^3$ )	$E_5$ ou $E_6$ .	$L_5$
. - de $D^4$ )	$E_3$ ou $E_4$ ou $E_5$ .	$L_6$
. - de $D^5$ )	$E_1$ ou $E_4$ .	$L_7$
. $D^3$ de $D^4$ )	$E_3$ ou $E_4$ ou $E_6$ ) ET	$L_8$
ET ( - de $D^5$ )	$E_1$ ou $E_4$ ou $E_5$ ou $E_6$ ) ET	$L_9$
ET ( $D^4$ de $D^5$ )	$E_1$ ou $E_3$ ou $E_5$ )	$L_{10}$

Dans un deuxième temps, nous rechercherons les sous-ensembles filtrants.

b.1. Méthode par énumération.

Il est aisé de voir que les sous-ensembles d'examens d'un ou de deux éléments sont insuffisants pour aboutir à un diagnostic valable. En effet, quatre combinaisons sont réalisables avec deux examens, pour cinq diagnostics possibles. Au minimum, trois examens seront donc nécessaires

pour discriminer les diagnostics.

D'une façon générale, le nombre minimum d'examens  $p$  nécessaires pour différencier  $n$  diagnostics doit être tel que

$$2^p \geq n$$

La méthode par énumération est très simple. Elle consiste à énumérer tous les sous-ensembles d'examens comprenant au minimum  $p$  éléments et de voir s'ils sont filtrants ou non. Rappelons qu'un sous-ensemble d'examens est filtrant s'il est discriminant pour chaque couple de diagnostics, c'est-à-dire lorsqu'il y a au moins un élément commun avec chacun des sous-ensembles du tableau des ET.

Le tableau suivant présente les sous-ensembles filtrants obtenus :

N°	SOUS-ENSEMBLES FILTRANTS	N°	SOUS-ENSEMBLES FILTRANTS
1	* E <sub>2</sub> E <sub>4</sub> E <sub>5</sub>	9	E <sub>2</sub> E <sub>4</sub> E <sub>5</sub> E <sub>6</sub>
2	* E <sub>1</sub> E <sub>2</sub> E <sub>3</sub> E <sub>5</sub>	10	E <sub>1</sub> E <sub>2</sub> E <sub>3</sub> E <sub>4</sub> E <sub>5</sub>
3	* E <sub>1</sub> E <sub>2</sub> E <sub>3</sub> E <sub>6</sub>	11	E <sub>1</sub> E <sub>2</sub> E <sub>3</sub> E <sub>4</sub> E <sub>6</sub>
4	E <sub>1</sub> E <sub>2</sub> E <sub>4</sub> E <sub>5</sub>	12	E <sub>1</sub> E <sub>2</sub> E <sub>3</sub> E <sub>5</sub> E <sub>6</sub>
5	* E <sub>1</sub> E <sub>2</sub> E <sub>4</sub> E <sub>6</sub>	13	E <sub>1</sub> E <sub>2</sub> E <sub>4</sub> E <sub>5</sub> E <sub>6</sub>
6	* E <sub>1</sub> E <sub>2</sub> E <sub>5</sub> E <sub>6</sub>	14	E <sub>2</sub> E <sub>3</sub> E <sub>4</sub> E <sub>5</sub> E <sub>6</sub>
7	E <sub>2</sub> E <sub>3</sub> E <sub>4</sub> E <sub>5</sub>	15	E <sub>1</sub> E <sub>2</sub> E <sub>3</sub> E <sub>4</sub> E <sub>5</sub> E <sub>6</sub>
8	* E <sub>2</sub> E <sub>3</sub> E <sub>4</sub> E <sub>6</sub>		

Les sous-ensembles marqués \* sont dits 'de base', ce qui signifie qu'aucun autre sous-ensemble filtrant ne leur est inclus.

Pour tenir compte des critères, il suffit de pondérer tous les sous-ensembles filtrants de base par la somme des coûts et des durées et de choisir celui qui optimise le critère considéré.

Cependant, lorsque le nombre d'examens et de diagnostics augmente, cette méthode devient lente et fastidieuse.

C'est pourquoi les méthodes graphiques ont été introduites.



### b.2. Méthode graphique d'exploration directe.

Cette méthode nécessite également l'élaboration préliminaire du tableau des ET, que l'on restructure en ordonnant ses lignes par nombre croissant d'éléments. Cette restructuration accélérera l'obtention des sous-ensembles filtrants.

Une deuxième manipulation de ce tableau consiste à restructurer chaque ligne du tableau ordonné en plaçant les examens dans l'ordre résultant du choix d'un critère.

Dans le tableau présenté ci-dessous, nous n'avons tenu compte que du critère coût, l'ordre utilisé étant celui des coûts croissants.

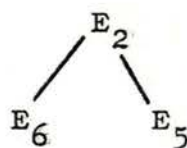
$T_1$	$E_2$
$T_2$	$E_6$ OU $E_5$
$T_3$	$E_4$ OU $E_1$
$T_4$	$E_6$ OU $E_2$ OU $E_5$
$T_5$	$E_4$ OU $E_3$ OU $E_5$
$T_6$	$E_6$ OU $E_4$ OU $E_3$
$T_7$	$E_1$ OU $E_3$ OU $E_5$
$T_8$	$E_6$ OU $E_2$ OU $E_4$ OU $E_3$
$T_9$	$E_6$ OU $E_4$ OU $E_1$ OU $E_5$
$T_{10}$	$E_6$ OU $E_2$ OU $E_4$ OU $E_1$ OU $E_5$

Ce tableau sert de base à la construction d'une arborescence où l'on adopte la règle 'toujours à gauche'.

Suivons les différentes étapes de la construction de cette arborescence.

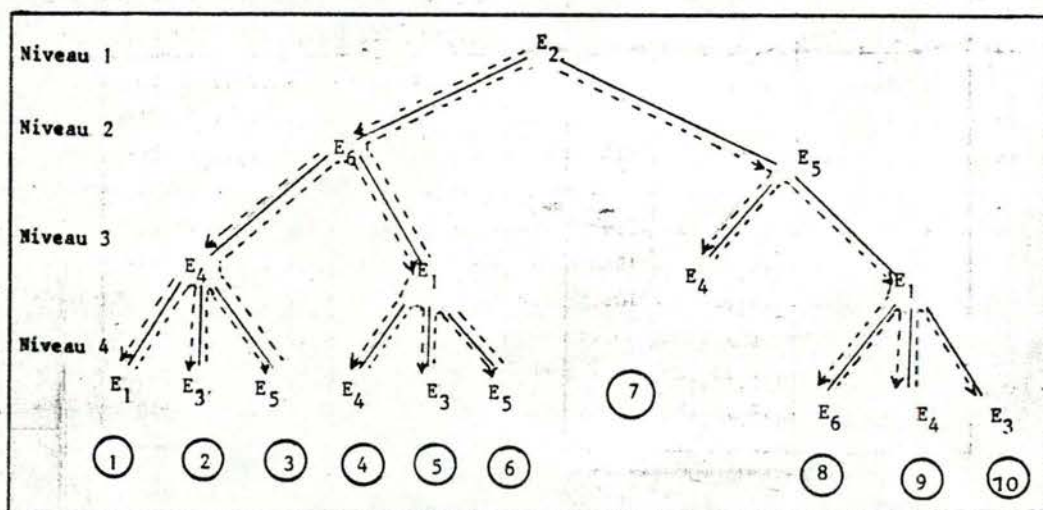
- étape 1 : On met la première ligne du dernier tableau présenté à la tête de l'arborescence. La racine est donc, dans notre exemple,  $E_2$ .
- étape 2 : On construit le second niveau avec les éléments de la deuxième ligne du même tableau à la condition qu'aucun élément de cette ligne ne figure au niveau supérieur.  
On obtient pour notre exemple le graphe

suisant :



- étape 1 : On construit le niveau suivant avec les éléments de la i-ième ligne du tableau. Ces éléments sont ajoutés à chaque branche de l'arborescence à la condition qu'aucun élément de la i-ième ligne ne figure pas dans un niveau supérieur.

Notre exemple aboutit à l'élaboration du graphe suivant :



Cette méthode permet d'exhiber rapidement une ou plusieurs chaînes filtrantes. En effet, une fois l'arborescence construite, la lecture 'toujours à gauche' nous permet d'obtenir un dernier tableau qui reprend, en regard des sous-ensembles filtrants obtenus, le poids du critère coût.

EXPLORATION DIRECTE AVEC LE CRITERE COUT		
Ordre d'obtention des chaines	sous-ensembles filtrants	coût (en B)
1	$E_1, E_2, E_4, E_6$	79
2	$E_2, E_3, E_4, E_6$	89
3	$E_2, E_4, E_5, E_6$	114
4	<i>idem</i> (1)	
5	$E_1, E_2, E_3, E_6$	113 *
6	$E_1, E_2, E_5, E_6$	138
7	$E_2, E_4, E_5$	106 *
8	<i>idem</i> (6)	
9	$E_1, E_2, E_4, E_5$	146
10	$E_1, E_2, E_3, E_5$	180



Cette méthode de construction est une adaptation de la procédure LIFO (Last In First Out) de la méthode du 'Branch and Bound' ou procédure S.E.P. (procédure par Séparation et Evaluation Progressive). Le dernier tableau montre que les premières chaînes obtenues sont celles qui minimisent le facteur coût. On constate cependant que l'ordre d'obtention de ces chaînes et la progression de leur poids ne sont pas strictement parallèles (voir \*).

Si, dans cet exemple, les premières chaînes exhibées sont celles qui coûtent le moins cher, il ne peut en être toujours de même. Cette méthode d'exploration directe est heuristique.

En résumé, elle guide le choix des examens complémentaires. Elle pourrait se coupler avec la méthode utilisée au CHU Timone de Marseille en ce qui concerne l'élaboration des fonctions employées lors du diagnostic proprement dit.

#### 2.1.1.B. Phase de classification.

C'est la phase du diagnostic.

Si le malade présente une symptomatologie qui donne à la fonction  $F_k$  la valeur vraie, alors on est sûr qu'il a la maladie  $k$ . Par contre, si toutes les fonctions  $F_k$  donnent la valeur fausse pour un pas donné, il est impossible de classer le malade car les informations connues à son sujet ne sont provisoirement pas suffisantes.

Il existe donc, non pas une branche d'erreur, mais une branche de rejet. Cette branche de rejet sera ensuite traitée de la même manière que les autres sous-branches d'affinement.

On pourrait définir cette méthode diagnostique pas à pas comme un enchaînement de décisions élémentaires auxquelles est soumis le patient et dont les éléments sont les fonctions logiques précédemment définies.

Comme on le voit, cette méthode suit exactement la démarche du médecin qui prend en compte, à chaque pas, une série

de nouveaux signes. A chaque itération, on incorpore ces signes dans les différentes fonctions logiques pour affiner le diagnostic.

### 2.1.2. Extension à l'algèbre ternaire.

\* principe :

Toute variable ne peut prendre que trois valeurs (1,0,-1) selon respectivement la présence, l'incertitude ou l'absence du signe dans la symptomatologie du patient considéré. Les relations entre symptômes et maladies s'expriment sous forme de relations logiques et peuvent se représenter sur une table de décision, après diverses simplifications analogues à celles opérées dans les autres méthodes logiques.

L'extension de l'algèbre de Boole à une algèbre à trois états a cependant nécessité la création d'opérateurs supplémentaires.

Voici une des façons de considérer le problème, celle adoptée par une équipe de recherches du Pr. Grémy (groupe de recherches U88-INSERM-PARIS).

+1 correspond à la présence.

0 correspond à l'incertitude.

-1 correspond à l'absence.

Le OU binaire devient le maximum.

Le ET binaire devient le minimum.

Le Complément devient la négation algébrique.

% signifie que l'incertitude devient absence.

@ correspond à la rotation positive.

Les opérateurs ternaires utilisés apparaissent sur les deux tables de vérité suivantes :



x	$\bar{x}$	$\oplus x$	$\%x$
1	-1	-1	1
0	0	1	-1
-1	1	0	-1

x	y	$x + y$	$x * y$
0	0	0	0
0	1	1	0
0	-1	0	-1
1	0	1	0
1	1	1	1
1	-1	1	-1
-1	0	0	-1
-1	1	1	-1
-1	-1	-1	-1

### 2.1.3. La démonstration automatique de théorèmes.

\* principe :

Cette méthode se fonde également sur les déductions logiques et elle peut intervenir à tous les pas diagnostiques.

La connaissance médicale relative à chaque diagnostic est mise sous forme d'axiomes. Ceux-ci constituent ce que l'on appelle, en logique du premier ordre, une théorie. L'association des signes présentés par le patient est considérée comme une formule de la théorie. Si cette formule peut se ramener à un des axiomes, alors le patient est affecté d'une maladie, celle dont il est question dans l'axiome.

\* critique des méthodes logiques.

Les méthodes logiques ne tiennent pas compte des arguments de fréquence. Elles se contentent d'exprimer la présence ou l'absence d'un symptôme sans tenir compte des fréquences d'apparition observées de cette existence ou de cette absence.

## 2.2. La méthode probabiliste (théorème de Bayes).

\* hypothèses de base :

- exclusivité et exhaustivité des diagnostics.

Pour être applicable, il faut qu'un patient donné ne soit atteint que d'une et une seule des maladies possibles, l'éventail de ces maladies devant être complet.

- indépendance des signes recherchés.

La réponse à un signe ne peut en aucun cas dépendre des réponses obtenues aux signes précédemment recherchés. Pareille hypothèse est cependant bien difficile à respecter au plan pratique. Aussi, certains auteurs ont essayé d'y porter remède en introduisant dans le processus bayésien la dépendance des signes,

- soit en cherchant à regrouper les signes deux à deux, en calculant le CHI-CARRE de toutes les paires possibles (Nugent),

- soit en substituant à l'hypothèse d'indépendance totale celle, non moins contraignante, de relations linéaires entre signes (Brunk, Lehr),

- soit en regroupant les signes exclusifs, de sorte qu'un seul signe de chaque groupe puisse être pris en compte dans le calcul du diagnostic (Warner, Toronto et Veasy),

- soit encore en établissant une formule qui tienne compte des corrélations de tout ordre entre les signes (Bahadur), ou en construisant un modèle original adapté aux signes dépendants.

Cependant, ces diverses tentatives ne restent applicables que pour un nombre de signes et de diagnostics relativement faible.

Ce problème de l'indépendance des signes est volontiers exagéré ou minimisé. A ce sujet, introduisons ici trois remarques :

- L'hypothèse d'indépendance des signes n'est



absolument pas nécessaire à l'utilisation d'un processus bayésien qui, sur le plan théorique, peut tout à fait s'en passer. Seules des considérations pratiques font que, le plus souvent, les expériences réalisées doivent assumer une telle hypothèse.

- Cette hypothèse est relative à chacun des diagnostics considérés. Cela revient à dire que la probabilité d'existence d'un signe quelconque, si tel diagnostic est présent, est la même si tel autre signe existe ou n'existe pas, ce qui ne traduit pas l'indépendance absolue entre signes :

$$P(S_i/S_j) = P(S_i)$$

Ce fait met en lumière l'un des intérêts de la première hypothèse, à savoir l'exclusivité des diagnostics, qui permet plus facilement d'analyser et éventuellement d'accepter certaines hypothèses d'indépendance.

- Bien souvent, les auteurs acceptent l'hypothèse de 'pseudo-indépendance' qui élimine uniquement les relations les plus fortes entre signes.

\* paramétrage :

La méthode probabiliste consiste à traduire un fichier de données, celui des dossiers médicaux, en probabilités, sur lesquelles se fonde l'application du théorème de Bayes.

Définissons les variables :

$ND_j$ , le nombre de malades de diagnostic  $D_j$ ,  
 $N_i$ , le nombre de malades qui présentent le signe  $S_i$ ,  
 $N_i^+$ , le nombre de malades qui ont le symptôme  $S_i$  positif,  
 $N_i^-$ , le nombre de malades qui ont le symptôme  $S_i$  négatif,  
 $N_{ij}$ , le nombre de malades de diagnostic  $D_j$  qui présentent le signe  $S_i$ ,  
 $N_{ij}^+$ , le nombre de malades de diagnostic  $D_j$  qui présentent le symptôme  $S_i$  positif,  
 $N_{ij}^-$ , le nombre de malades de diagnostic  $D_j$  qui présen-

tent le symptôme  $S_i$  négatif.

L'on peut donc estimer les paramètres suivants sur base d'un échantillon de malades de diagnostic confirmé.

- probabilité de chaque diagnostic.

$$P(D_j) = \frac{ND_j}{\sum_k ND_k}$$

où  $P(D_j)$  représente la probabilité a priori du diagnostic  $D_j$ .

- probabilité de chaque signe.

$$P(S_i) = \frac{N_i^+}{N_i}$$

$$Q(S_i) = \frac{N_i^-}{N_i} = 1 - P(S_i)$$

où  $P(S_i)$  représente la probabilité a priori que le

symptôme  $S_i$  apparaisse de manière positive.

- probabilité de chaque signe relative à chaque diagnostic.

$$P(S_i/D_j) = P(S_i^+/D_j) = \frac{N_{ij}^+}{N_{ij}}$$

$$Q(S_i/D_j) = \frac{N_{ij}^-}{N_{ij}}$$

où  $P(S_i/D_j)$  représente la probabilité conditionnelle de  $S_i$  par rapport à  $D_j$ .

Comme nous le verrons plus loin, l'estimation de ces différents paramètres pose bien souvent de nombreux problèmes pratiques.



\* calcul du diagnostic :

La connaissance d'un signe  $S_i$  permet d'ajuster les probabilités a priori des différents diagnostics grâce à l'application pure et simple du théorème de Bayes :

$$P(D_j/S_i) = \frac{P^o(D_j) * P(S_i/D_j)}{\sum_k P^o(D_k) * P(S_i/D_k)}$$

où  $P^o(D_j)$  représente la probabilité du diagnostic  $D_j$  avant la prise en compte de l'observation du signe  $S_i$ . Initialement, lorsqu'aucun signe n'est encore considéré,  $P^o(D_j)$  représente la probabilité a priori du diagnostic  $D_j$ .

L'hypothèse d'indépendance des symptômes permet d'affirmer, par exemple, que

$$P(D_j/S_1^+, S_2^-, S_3^+) = \frac{P^o(D_j) * P(S_1/D_j) * Q(S_2/D_j) * P(S_3/D_j)}{\sum_k P^o(D_k) * P(S_1/D_k) * Q(S_2/D_k) * P(S_3/D_k)}$$

De plus, l'absence d'un symptôme peut présenter autant d'intérêt que la présence d'un autre et le théorème de Bayes permet précisément d'en tenir compte.

La prise en considération des signes observés chez le patient modifie la valeur du vecteur des probabilités des diagnostics en probabilités a posteriori. Pour interpréter ces valeurs de probabilités, nous proposons trois critères différents :

- considérer l'ordre (probabilité maximum = diagnostic réel),
- établir un seuil de décidabilité qui fixe un diagnostic en cas de dépassement de ce seuil, qui propose des examens complémentaires ou qui

laisse planer le doute,

- considérer la cardinalité des probabilités des diagnostics, dans l'espoir d'y trouver des différences significatives entre les premiers classés.

Quant à la détermination d'un diagnostic différentiel, nous renvoyons le lecteur au paragraphe relatif au choix des examens complémentaires.

\* critique :

L'estimation des probabilités a priori et des probabilités conditionnelles nécessite des statistiques convenablement rassemblées et mises à jour, ce qui n'est pas toujours le cas dans certains milieux médicaux.

Lorsque l'échantillon se révèle trop petit, on fait appel à des probabilités 'subjectives' déterminées par des médecins spécialistes et progressivement revues et affinées. L'expérience a montré que ces évaluations cernaient la réalité avec beaucoup de précision, mis à part quelques cas particuliers.

L'indépendance des symptômes entrant dans les différents profils des patients n'est pas toujours vérifiée, loin s'en faut. Cependant, cette lacune n'a pas nui à l'efficacité pratique des différents programmes d'aide au diagnostic actuellement mis au point.

Ainsi que nous l'avons vu, certains auteurs ont essayé d'y remédier sans obtenir pour autant des résultats sensiblement meilleurs.

L'existence de signes ou de syndromes qui excluent ou qui identifient définitivement un diagnostic ou un groupe de diagnostics n'est pas considérée.

Enfin, pour l'élaboration du diagnostic, seules les données qualitatives peuvent être directement utilisées si bien que les données quantitatives, réduites à des valeurs qualitatives (intervalles de valeurs, groupages,...) perdent un peu de leur valeur informationnelle.



### 2.3. La théorie des sous-ensembles flous.

\* principe :

Tout développement mathématique doit suivre la perception humaine. S'il est impossible d'affecter des valeurs précises dans notre pensée, notre perception doit toujours être exacte et fidèle. Faute pour l'imprécis mais rejetant l'inexactitude, telle apparaît la théorie des sous-ensembles flous. C'est peut-être parce que tant de gens ont confondu l'imprécis et l'inexact qu'une injuste méfiance a prévalu pendant plusieurs années vis-à-vis de cette théorie.

La théorie des sous-ensembles flous est une théorie des connaissances imprécises et qui ne peut accepter les raisonnements inexacts; elle se propose de nous faire utiliser des modèles où l'on peut accepter des données, des relations, des procédures, des structures qui sont floues, c'est-à-dire qui sont perçues et utilisées avec leur imprécision naturelle, mais dans le but d'obtenir des décisions qui concordent au mieux avec les intentions des décideurs. Cette théorie est fondée sur la notion d'appartenance. Mais on ne se limite pas aux deux phénomènes classiques d'appartenance ou de non-appartenance.

Si l'on se place dans un contexte médical avec un ensemble de patients, peut-on dire d'une personne : elle appartient ou elle n'appartient pas à l'ensemble des maladies ? Non, car les limites d'une telle affirmation sont floues. C'est pour cette raison que l'on a introduit la notion de 'degré d'appartenance'.

On voit très bien l'utilité de cet outil lorsque l'on travaille sur des critères tels que :

- une personne a-t-elle de l'humour ?
- une personne est-elle irritable, jolie ou intelligente ?
- une personne est-elle atteinte de telle maladie mentale ?

Ainsi, les réponses à de telles questions et l'établissement de pareils diagnostics entraînent des constructions abstraites qui constituent les sous-ensembles flous, c'est-à-dire des classes aux limites imprécises dans lesquelles la transition d'appartenance est plus graduelle que brusque.

Cette théorie développée par le Pr. L.A. Zadeh de l'université de Berkeley a été introduite en France par le Pr. A. Kaufmann.

Un sous-ensemble flou se définit par la liste des différents éléments auxquels est attribué un degré d'appartenance à un ensemble de référence et non, comme pour les ensembles traditionnels, par l'énumération des objets qui le composent.

Dès à présent, nous pouvons dire que l'idée qui est à la base de cette théorie s'adapte bien au domaine médical. En effet, le médecin est souvent confronté à des données, des appréciations imparfaitement définies et il est amené à interpréter ces informations en tant que telles dans leur propre domaine non mesurable. De plus, il existe souvent des trous d'information.

Cette théorie permet aussi de prendre en compte l'importance qu'il convient d'attribuer à l'absence d'un symptôme qui, pour diverses raisons, n'a pas été recherché.

Une équipe de Marseille, sous l'impulsion de R. Sambuc, a généralisé la notion de sous-ensemble flou en introduisant le concept de sous-ensemble  $\tilde{F}$ -flou qui permet de traduire à la fois la notion de flou qui caractérise l'appartenance pondérée à un ensemble et la notion d'indétermination qui peut s'attacher à l'évaluation même de cette appartenance.



## B. Les méthodes de choix d'examens complémentaires.

L'utilisation des méthodes d'aide au diagnostic s'avère plus féconde lorsqu'elle permet l'élaboration d'une stratégie diagnostique par étapes successives. A chaque itération, le programme énonce les diagnostics encore envisageables assortis de leurs probabilités respectives. La stratégie sera efficace si elle aboutit, en réduisant progressivement le nombre des options, à augmenter la probabilité de celles qui restent en piste. Idéalement, il faudrait arriver à un diagnostic affecté de la probabilité 1.0 (certitude).

Nous pouvons représenter l'ensemble des possibilités diagnostiques sous la forme d'un graphe arborescent. Chaque noeud de l'arbre symbolise la recherche d'un test clinique ou biologique et chaque branche, l'issue possible d'un test qui conduit à un test ultérieur. La détermination de la stratégie optimale se ramène au problème de la recherche du chemin de longueur minimale dans un graphe séquentiel.

Comme nous allons le montrer, la théorie de la décision permet de résoudre le problème de la détermination d'une telle stratégie optimale d'investigations. Elle implique que l'on sache définir une fonction d'utilité, c'est-à-dire associer un coût global à chacune des investigations et une utilité à chacune de leurs conséquences diagnostiques ou thérapeutiques.

Nous développerons ensuite un modèle emprunté à la théorie de l'information, avant de conclure par la présentation d'une étude originale et intéressante faite à ce propos.

### §1. Théorie de la décision.

\* méthode : algorithme de Ford

Soient  $D_1, D_2, \dots, D_n$ , les  $n$  diagnostics exhaustifs et exclusifs de probabilités a posteriori  $P(D_1), P(D_2), \dots, P(D_n)$  calculées à l'aide du théorème de Bayes.

Soient  $E_1, E_2, \dots, E_r$ , les examens complémentaires dont le résultat ne peut être que positif ou négatif, un examen positif permettant d'affirmer un diagnostic alors qu'un résultat négatif ne permet pas de conclure avec précision.

Soit  $q_d(i_1, i_2, \dots)$ , la probabilité que  $i_1, i_2, \dots$  aient été négatifs pour le diagnostic  $d$ .

Soit  $C_d(i_i)$ , le coût de l'examen  $i$  pour le diagnostic  $d$ . Le coût d'une erreur de diagnostic est considérable vis-à-vis du coût relatif de chaque examen si bien que l'on ne cesse les investigations tant que l'on n'obtient pas une réponse positive.

Le seul problème qui se pose alors est celui de l'ordre des examens.

Si nous considérons une suite d'examens  $E_{i1}, E_{i2}, \dots, E_{ij}$ , son coût total est égal à la somme des coûts des examens qui la composent. Le coût moyen de  $E_{ij}$  est :

$$C_d^0(E_{ij}/E_{i1}, E_{i2}, \dots) = \sum_d P(D_d) * C_d(ij) * q_d(i1, i2, \dots)$$

puisque  $E_{ij}$  n'est effectué que si tous les examens précédents ont tous été négatifs. Le coût de l'examen  $E_{ij}$  ne dépend en fait que des examens qui le précèdent et non de leur ordre, les résultats de ces examens étant indépendants.

Nous en sommes ainsi ramené au problème classique de la recherche du chemin de longueur minimum dans un graphe séquentiel, résolu aisément par l'application de l'algorithme de Ford.

\* critique :

L'algorithme de Ford entraîne une masse de calculs qui le rend inapplicable lorsque le nombre d'examens s'accroît.

Le coût de chaque examen est une variable agrégée du coût



financier de l'examen, du délai d'obtention de son résultat et du risque de complication que peut entraîner l'attente de ce résultat. Cependant, il faudrait pouvoir tenir compte d'une foule d'autres facteurs dont certains varient en chaque instant, telle la disponibilité du médecin devant pratiquer l'examen ou encore l'encombrement actuel du laboratoire.

L'évaluation subjective, car leur prise en considération dépend de chaque médecin, d'une fonction de coût unique qui intégrerait des facteurs si divers et si évolutifs pose un problème très ardu, dont la solution est, par là même, difficilement formalisable.

Enfin, l'algorithme de Ford, tel qu'il a été implémenté par Lellouch J. (B42) dans une application d'aide au diagnostic reste acceptable si et seulement si le nombre d'examens complémentaires proposés ne dépasse pas la quinzaine, s'il en existe au moins un qui fournisse à coup sûr le diagnostic, ce qui n'est évidemment pas le cas le plus courant en matière d'examens complémentaires.

Dans de nombreux cas pratiques en effet, le diagnostic ne peut être obtenu, quels que soient les examens réalisables. Compte tenu de cette réalité, le médecin peut être amené à interrompre ses investigations sans avoir obtenu de certitude. Cette décision, qui relève de sa seule et unique responsabilité, repose sur la comparaison des coûts et de l'intérêt des examens restant à réaliser avec l'inconvénient que présente le fait de ne pas conclure.

Certains ont d'ailleurs tenté d'introduire dans le processus un coût de non-conclusion.

## §2. Théorie de l'information.

\* méthode :

A l'instar de l'algorithme de Ford, cette méthode propose un processus de recherche pas à pas, que nous pourrions qualifier de sub-optimal.

Si le vecteur des diagnostics  $D_1, D_2, \dots, D_n$  admet pour probabilités  $P(D_1), P(D_2), \dots, P(D_n)$  à une étape donnée, une mesure de l'indétermination qui subsiste quant à l'incertitude du diagnostic en cause peut se faire au moyen de la fonction entropie  $H(T)$ .

$$H(T) = \sum_{j=1}^n P(D_j) * \log \frac{1}{P(D_j)}$$

Si l'on opère un test d'issues possibles  $S_1, S_2, \dots, S_a$ , les probabilités des diagnostics évolueront différemment en chacune de ces issues et l'entropie du système en sera elle aussi modifiée.

Soit alors  $H'(T)$ , l'entropie du système  $T$  après le test  $q$ , la moyenne étant opérée sur l'ensemble des issues de ce test. La différence  $D$  :

$$D = H(T) - H'(T)$$

mesure dès lors l'information qu'apporte le test  $q$  considéré, c'est-à-dire, en quelque sorte, son pouvoir diagnostique.

\* critique :

L'objection majeure à l'encontre de cette méthode met tout simplement en cause la mesure même que fait la fonction entropie  $H(T)$  de l'indétermination du vecteur des probabilités a posteriori.

D'autre part, pour ceux qui acceptent malgré tout la validité de cette mesure, il faut reconnaître que le choix des tests complémentaires devrait dépendre, non seulement du gain d'information apporté par le test, mais également d'autres facteurs tels que la sécurité du malade, le risque qu'il encourt, l'attente ...etc.

Il faudrait donc, comme dans la première optique, définir une fonction de coût globale.



### §3. Heuristiques.

Nous devons citer ici une étude fort intéressante réalisée par Le Minor, Alpérovitch et Lellouch (réf.) et qui nous propose deux méthodes inspirées de la théorie de la décision.

#### 3.1. Méthode S.

Cette heuristique consiste à limiter la profondeur de l'exploration de l'arbre de décision à  $k$  examens (méthode  $S_k$ ).

La stratégie proposée est alors la meilleure des

$$\sum_{p=1}^k A_n^p \quad \text{stratégies de longueur inférieure}$$

ou égale à  $k$ , parmi  $n$  examens possibles.

En cas de non-conclusion, il suffit de répéter la même méthode de sélection d'examens, autant de fois qu'il sera nécessaire.

#### 3.2. Méthode $\Sigma$ .

Chaque fois qu'il doit déterminer le test à réaliser, le médecin écarte, en fonction de ses hypothèses de diagnostic, celles dont la valeur diagnostique lui paraît trop faible ou le risque encouru trop élevé, et opère son choix sur les tests qui restent pertinents après ce filtre. Les investigations rejetées ne le sont pas définitivement car un choix peut toujours être révisé après chaque résultat.

---

Le Minor M., Alpérovitch A., Lellouch J.

'Prescription séquentielle des examens complémentaires en médecine. Etude par simulation de quelques heuristiques.'

### 3.3. Méthode S.

Combinaison des deux heuristiques précédentes, cette optique limite l'exploration de l'arbre de décision en profondeur ( $k$  examens) et au niveau de chaque noeud ( $j$  branches).

### 3.4. Pas à pas.

Vu sous cet angle, le pas à pas, tel qu'il a été considéré précédemment, devient un cas particulier des méthodes précédentes, à savoir  $\Sigma_1$ , ou une répétition de  $S_1$  ou encore de  $\Sigma_1 S_1$ .

Les auteurs précités ont réalisé une étude par simulation de ces différentes heuristiques. Ils ont appliqué une forme de l'algorithme de Ford aux différentes heuristiques qu'ils nous ont proposées.

Au delà des chiffres qu'ils présentent, retenons trois de leurs conclusions :

- La prescription pas à pas se révèle être sensiblement la plus mauvaise des méthodes.

- Plus la stratégie optimale est longue, plus la limitation de l'exploration de l'arbre est pénalisante.

- D'un point de vue pratique, la stratégie optimale s'exclut généralement à cause du nombre d'examens envisagés.

Le médecin peut proposer des valeurs de  $k$  (méthode  $S_k$ ) et de  $j$  (méthode  $\Sigma_j$ ) raisonnables, qu'il suffit alors d'introduire dans le processus de simulation afin de juger de l'opportunité du choix d'une heuristique.



## CHAPITRE 4. DEUX EXEMPLES D'APPLICATION.

### A. Application des méthodes d'analyse multidimensionnelle au pronostic de l'infarctus myocardique.

(J.P. NAKACHE, P. LORENTE)

(Groupes de recherches INSERM U88 et U14, Paris)

#### §1. Le problème thérapeutique.

Le malade atteint d'un infarctus myocardique aigu compliqué d'une défaillance de la pompe cardiaque peut bénéficier de certaines techniques d'assistance parmi lesquelles la plus utilisée est la contre-pulsion par ballon intra-aortique. Cette thérapeutique nécessite l'intervention d'une équipe médico-chirurgicale entraînée et d'un matériel important, qui ne se justifie que si les chances de survie paraissent réduites.

#### §2. Les variables utilisées dans l'étude.

##### 2.1. Les patients.

L'étude est basée sur 101 patients présentant un infarctus myocardique aigu compliqué de défaillance cardiaque, sans aucune autre affection, parmi lesquels figurent 52 infarctus du myocarde antérieur et 49 infarctus du myocarde postérieur. Est considéré comme décédé, le malade pour lequel une issue fatale survient dans un délai de 4 semaines après le début de l'infarctus.

##### 2.2. Les variables.

Elles sont présentées sur le tableau 1 de l'annexe 1 :

- 19 variables quantitatives,
- 3 questions à deux modalités (A),

- la variable à expliquer (V),
- la localisation de l'infarctus (antérieur, postérieur) (L), variable qui apparaîtra d'elle-même sur le deuxième axe factoriel.

§3. Etude de l'échantillon de base par l'Analyse des Correspondances.

Voici les pourcentages d'inertie expliqués par les 5 premiers facteurs de l'AFC (calculés à partir du tableau de Burt).

<u>Axes_factoriels</u>	<u>Pourcentages</u>	<u>Pourcentages_cumulés</u>
F1	35.636	35.636
F2	14.350	49.986
F3	10.860	60.846
F4	5.507	66.353
F5	4.744	71.097

La figure 1 ci-dessous présente la projection sur le plan (F1, F2) des 101 patients et des paramètres. L'examen de la figure montre une séparation entre les survivants et les décédés (dans 4 semaines), de façon très nette d'ailleurs pour les infarctus myocardiques antérieurs (A).

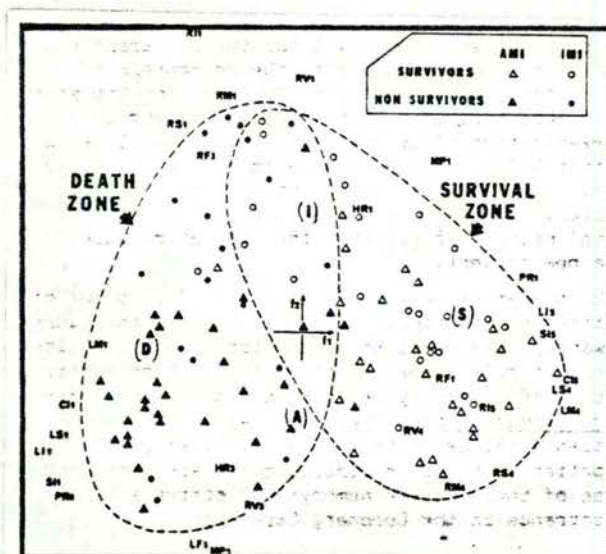


Figure 1 : Représentation des 101 patients et des paramètres sur le plan (F<sub>1</sub>, F<sub>2</sub>).



Dans la zone de survie, on dénombre 43 survivants et 2 décédés.

Dans la zone de décès, on compte 42 décédés et 2 survivants.

Dans l'intersection, figurent 5 survivants et 7 décédés.

Au total, 86% de survivants sont bien classés, 82% de décédés le sont également.

Sur l'ensemble des 101 patients, 12 (soit 12%) sont mal situés.

#### §4. Pronostic et décision thérapeutique.

##### 4.1. Choix d'une méthode de discrimination.

Devant un cas nouveau d'infarctus aigu compliqué de défaillance cardiaque, il s'agit de reconnaître rapidement si le patient n'a que de faibles chances de survie en vue d'une assistance circulatoire éventuelle.

Le processus comprend deux étapes :

- Adjonction du cas nouveau comme élément supplémentaire s aux résultats de l'analyse des 101 dossiers de base. Le calcul des facteurs  $F_a(s)$  se fait par la formule:

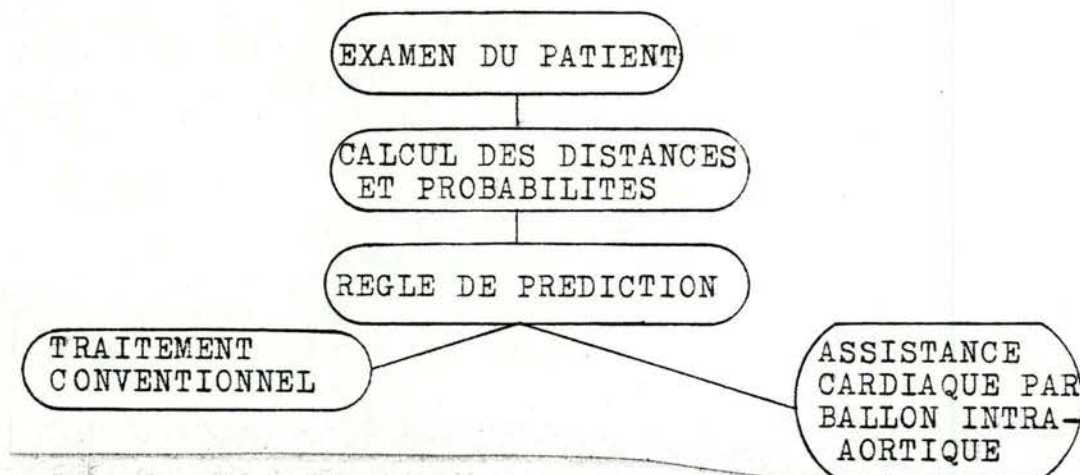
$$F_a(s) = (\lambda_a^{-1/2} / k(s)) * \sum \{ k(s, j) * G_a(j) \mid j \in J \}$$

où  $k(s, j) = 1$  si  $s$  possède la modalité  $j$ ,  
 $= 0$  sinon.

- Prise en considération du point  $S = (F1(s), F2(s))$  dans le plan  $(F1, F2)$  et interprétation de sa position par rapport à ses proches voisins.

##### 4.2. Application à 55 cas nouveaux.

Voici le diagramme fonctionnel du pronostic de chaque nouveau cas considéré :



Pour juger de l'efficacité de la méthode, il suffit de considérer la figure 2 ci-dessous qui situe sur les axes F1 et F2, les 55 cas nouvellement considérés.

Parmi eux, 21 cas se situent franchement dans la zone de survie. Le traitement médical classique leur est appliqué et tous survivent.

34 cas se situent dans ou à proximité de la zone de décès. 25 de ces patients peuvent bénéficier de l'assistance circulatoire. 9 d'entre eux passent le cap des 4 semaines, les 16 autres décèdent.

Tous ceux qui n'ont pu profiter de l'assistance décèdent également.

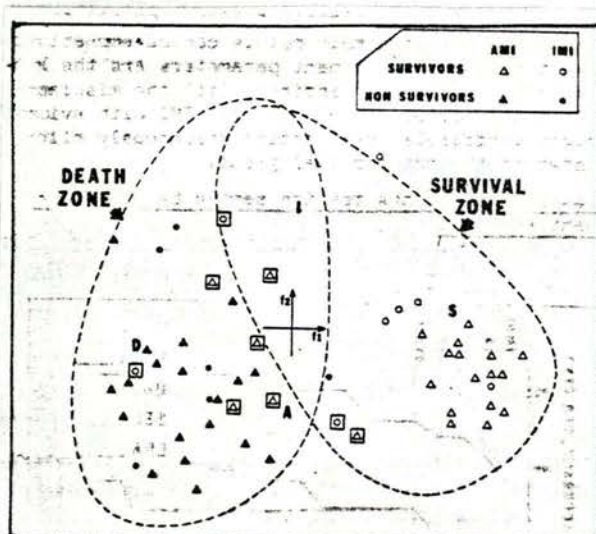


Figure 2 : Application à 55 cas nouveaux.



## §5. Conclusion.

- Confirmation de l'efficacité de la technique d'assistance de contre-pulsion par ballon intra-aortique, bien que 5 des 9 patients ayant survécu 4 semaines soient décédés dans une période de un à plusieurs mois. "Cette mortalité différée paraît due, souligne l'étude, à un état général très altéré du réseau coronarien, au-dessus de tout recours chirurgical."

- Démonstration de l'aide présentée au praticien par les méthodes multidimensionnelles dans le pronostic et la décision thérapeutique. Par des calculs faciles, elles permettent la confrontation d'un cas nouveau à une représentation géométrique synthétisant l'analyse des cas antérieurs et aisément interprétable.

## B. Application de la méthode probabiliste au diagnostic en neurologie.

(Groupe de recherches U88-INSERM, Paris)

### §1. Origine.

L'objectif principal du système mis en place est pédagogique. Il vise à proposer une simplification du contenu de l'enseignement de la neurologie jugé trop complexe et mal adapté aux véritables besoins du futur généraliste.

La neurologie peut sans doute se caractériser d'autres domaines médicaux sous certains aspects qui matérialisent précisément la difficulté de son apprentissage et des problèmes d'interprétation qu'elle pose :

- sa sémiologie très vaste et très complexe,
- diagnostics très divers,
- nécessité d'un raisonnement logique,
- coût élevé des examens complémentaires, donc importance de la pertinence du choix.

Ces différents points démontrent l'adéquation de cette pathologie à certains avantages offerts par l'automatisation de l'aide au diagnostic.

## §2. Source des données.

Deux neurologues participent à l'expérience avec des hypothèses et des dossiers propres.

- Le professeur SAMSON de Rouen (service de neurologie, hôpital Charles Nicolle).

Son étude porte sur la démarche diagnostique en neurologie. Il analyse 21 diagnostics topographiques décrits par 99 signes répartis en 14 syndromes.

- Le professeur DEROUESNE de Paris (clinique des maladies du système nerveux, la Salpêtrière).

Son objectif vise à simplifier le contenu et la méthode de l'enseignement de la neurologie. L'hypothèse de travail qu'il développe porte sur 31 diagnostics (la sélection faite est précisément à améliorer) et sur 166 signes.

## §3. Saisie des données.

Actuellement, près de 1000 dossiers constituent la base de données exploitable. Pour chaque dossier, le diagnostic réel a pu être confirmé et l'on n'a dès lors tenu compte que des signes constatés lors du premier examen neurologique de ces malades, c'est-à-dire avant que des examens complémentaires neurochirurgicaux aient été pratiqués.

Les protocoles ont été et sont encore recueillis par des internes ou des étudiants de C.E.S. intégrés à l'ensemble de l'étude et, de ce fait, parfaitement conscients de sa portée.

## §4. L'outil informatique.

### 4.1. Hardware.

- PDP 10 du CHU Pitié-Salpêtrière.



## 4.2. Software.

### 4.2.1. Méthode utilisée pour le calcul du diagnostic.

- Utilisation de la méthode probabiliste (Bayes).
- Hypothèses classiques d'exclusivité des diagnostics et d'indépendance des symptômes.

### 4.2.2. Les probabilités à priori.

Dans l'optique pédagogique fixée, il a été jugé préférable d'initialiser le processus bayésien par une équiprobabilité des différents diagnostics, tout en conservant la possibilité d'incorporer, en dernier ressort, la notion de fréquence à priori.

### 4.2.3. Les probabilités conditionnelles.

Se trouvant dans l'impossibilité de recueillir un échantillon de dossiers suffisamment représentatif de l'ensemble des diagnostics, les concepteurs ont eu recours aux évaluations des spécialistes. Qui dit évaluation, dit subjectivité, et c'est précisément pour corriger quelque peu ces probabilités subjectives que l'on a fait appel, dans la mesure du possible, aux fréquences observées, selon la formule :

$$P(S/D) = a * F(S/D) + (1-a) * P(x)$$

dans laquelle  $P(x)$  est la probabilité estimée,

$F(S/D)$  est la fréquence du signe  $S$  dans le diagnostic  $D$ , déterminée sur l'effectif  $E(D)$ ,  
 $a$  tel que  $0 < a < 1$  est choisi  $= \frac{E(D)}{K + E(D)}$ ,

$K$  étant une constante dont la valeur, après essais, a été fixée à 10.

Comme nous pouvons le constater, c'est une situation de compromis basée sur l'expérience des spécialistes et constamment affinée au fur et à mesure de l'introduction de nouveaux dossiers dans la base de données.

#### 4.2.4. L'indépendance des signes retenus.

C'est la pierre d'achoppement de l'ensemble des applications exploitant la méthode bayésienne. L'indépendance des signes (ou la prise en compte des interactions) pose de multiples problèmes à tous ceux qui, jusqu'aujourd'hui, ont appliqué cette méthode.

Cependant, différentes améliorations non négligeables permettent d'approcher cette notion d'indépendance :

- regroupement de signes exclusifs en un seul signe,
- élimination de signes redondants,
- regroupement de signes non-indépendants entre eux.

A l'intérieur de ces groupes, une priorité a été affectée, de sorte qu'un seul signe du groupe est pris en compte même si plusieurs sont présents, sans pour autant considérer les autres signes comme absents. Ces relations de priorité sont traduites par des expressions logiques au niveau de l'interface nécessaire entre le recueil des signes et leur incorporation dans le calcul du diagnostic. Les signes utilisés étant des variables logiques ternaires (présence, incertitude, absence), il a fallu élargir les opérateurs binaires classiques. (cfr. chapitre 3, point A/§2/2.1.2.)

#### 4.2.5. Diagnostic utile.

Il arrive fréquemment que l'un des diagnostics proposés sorte de la liste des probabilités de façon indiscutable, permettant de croire, avec un degré raisonnable de certitude qu'il s'agit là du diagnostic exact. Mais il peut également arriver qu'aucun diagnostic n'émane distinctement des autres. C'est à cette fin que les concepteurs du système ont introduit la notion du diagnostic utile, pour déterminer le diagnostic pour lequel le praticien a intérêt à agir.

Les concepteurs du projet ont défini une matrice d'utilité  $U$  où  $U(i,j)$  représente l'utilité de l'acte adapté au diagnostic  $D_i$ , lorsque le vrai diagnostic est  $D_j$ . Les valeurs ont, une fois encore, été estimées par les médecins.

Le diagnostic finalement retenu n'est pas nécessairement le



plus probable.

On choisit  $D_k$  tel que

$$U_k = \max_j U_j \quad \text{où } U_j = \sum_i P(D_i) * U(j, i)$$

où  $P(D_i)$  est la prob. a priori du diagnostic  $D_i$ .

Cette notion de diagnostic "utile" ou "opérationnel" n'a été appliquée qu'à la seconde hypothèse (DEROUESNE).

#### 4.2.6. Choix des examens complémentaires.

##### Hypothèses :

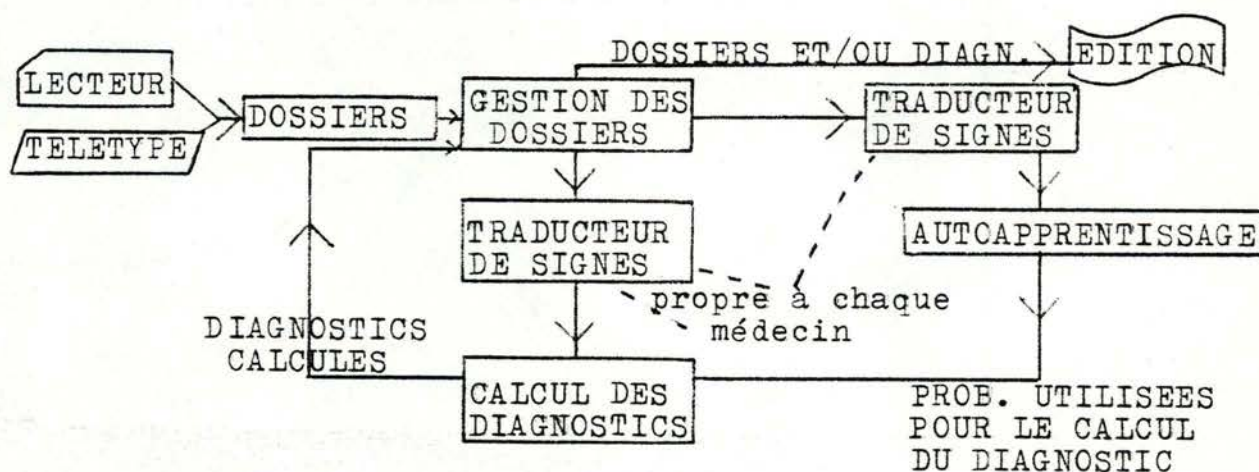
- Tous les examens ont un coût unique et identique.
- L'amélioration de l'utilité de la conclusion prime, aux yeux des concepteurs, sur l'optimisation du vecteur des probabilités des diagnostics, jugées purement descriptives, non opérationnelles.

##### Méthode adoptée :

- Soit  $D_k$ , le diagnostic présenté comme le plus utile après la prise en compte des signes cliniques.
- Si  $U_k$  dépasse un seuil fixé à l'avance, aucun examen complémentaire n'est jugé nécessaire.
- Si  $U_k$  n'atteint pas ce seuil, les tests sont sélectionnés de la manière suivante : doivent être prescrits tous les examens susceptibles - d'augmenter  $U_k$  pour le rendre supérieur au seuil fixé
  - ou de déterminer un diagnostic  $D_i$  avec  $i \neq k$ . A ce moment, il faut reposer la question initiale avec le nouveau vecteur des probabilités.

#### 4.2.7. Enchaînement des fonctions essentielles.

Voici le graphe des relations d'échanges d'informations.



On peut donc considérer deux sous-systèmes.

a. Gestion des dossiers.

- recueil et introduction des dossiers.
- restitution des dossiers en clair.
- correction des dossiers déjà introduits.
- sélection de dossiers sur critères variables ( recherche séquentielle non optimisée).

b. Calcul du diagnostic.

- introduction des matrices de probabilités initiales.
- définition des signes pris en compte pour le calcul du diagnostic et traduction au dossier.
- auto-apprentissage.
- calcul du diagnostic.
- évaluation globale et analyse détaillée des résultats.
- calcul et édition des résultats intermédiaires.

L'originalité principale de ce système réside dans la possi-



bilité de prendre en compte, pour le calcul du diagnostic, un ensemble de signes déduits logiquement des signes du dossier. Pour ce faire, un langage de définition des signes à partir de ceux du dossier a été créé. Il permet de définir chaque item retenu par le neurologue comme expression logique des signes du dossier.

## §5. Exploitation et utilisation.

A l'heure actuelle, le système mis en place s'exploite en une sorte de circuit fermé, c'est-à-dire que, sous l'hypothèse du professeur Derouesné, les 1000 dossiers mémorisés (en juin 78, date de fin de contrat de recherche) servent à la fois d'éléments de l'échantillon de base et de cas nouveaux étudiés. On reconsidère comme cas à analyser les patients qui ont servi à initialiser les données du calcul. Faute de crédits, toutes les investigations ou les enquêtes souhaitées se limitent à cet ensemble de dossiers. C'est évidemment un peu dommage dans la mesure où les résultats que l'on peut en tirer deviennent, par ce fait même, un peu biaisés.

## §6. Analyse des résultats.

### 6.1. Analyse globale.

#### 6.1.1. Professeur Samson.

Les premiers résultats obtenus furent décevants car pour 550 dossiers, 60% seulement furent bien classés (diagnostic exact proposé comme le plus probable). Cet échec pouvait s'expliquer par le fait que le processus utilisé supposait l'exclusivité des diagnostics alors que, dans le diagnostic topographique, les associations de foyers atteints sont fréquentes.

Une méthode permettant l'analyse détaillée des vecteurs de probabilités a posteriori a été adaptée au processus

bayésien. Pour cela, il a fallu :

- déterminer le caractère unilatéral ou bilatéral des lésions,
- établir une règle de décision permettant de retenir, parmi l'ensemble des diagnostics, les seuls dont la probabilité calculée les rendait vraisemblables (règle choisie : probabilité à posteriori  $>0.12$ ),
- définir des relations de "voisinage" entre les différents foyers afin de déterminer le caractère unique, multiple ou diffus de l'affection.

Le pourcentage de bons résultats (de plus forte probabilité) atteint alors 74% sur le même échantillon de dossiers avec une indication supplémentaire, la latéralité de l'affection, et dépasse 90% si l'on considère les deux premiers diagnostics proposés.

Cependant, et nous reprenons ici la critique même qui a été faite par les auteurs du système, "rien ne permet d'affirmer que l'erreur introduite par l'utilisation d'un modèle bayésien assumant l'exclusivité des diagnostics puisse être atténuée ou éliminée par une analyse, aussi fine soit-elle, du vecteur de probabilités à posteriori... En fait, tout le problème est de savoir si la symptomatologie qui résulte d'une atteinte de deux foyers neurologiques représente la "réunion" des signes induits par l'atteinte de chacun des foyers ou réalise un nouveau tableau clinique peu en rapport avec chacun des foyers touchés. La réponse à une telle question est essentiellement médicale et est délicate."

A l'heure actuelle, les auteurs se proposent d'utiliser une méthode logique pour le diagnostic topographique en neurologie.



### 6.1.2. Professeur\_Derouesné.

Les résultats apparaissent plus encourageants. Sur 550 dossiers analysés, 70% présentaient le diagnostic réel comme le plus probable, 82% comme le plus "utile" et 92% comme l'un des deux plus probables.

## 6.2. Analyse par diagnostic.

### 6.2.1. La matrice de coïncidence.

Dans cette matrice, le terme générique  $M(i,j)$  représente le nombre de malades de diagnostic réel  $D_i$  pour lesquels l'ordinateur a proposé  $D_j$ .

On dispose ainsi, pour chaque diagnostic considéré :

- du nombre de résultats exacts,
- du nombre de cas où tel autre diagnostic a été porté par erreur à la place de ce diagnostic,
- du nombre de fois où ce diagnostic a été proposé en lieu et place de tel autre.

Ces différents éléments présentent un intérêt évident pour le médecin lui-même, car ils répondent aux préoccupations que pose le diagnostic différentiel.

Ils mettent également en évidence certaines erreurs d'évaluation des probabilités subjectives.

Il serait cependant souhaitable que le terme générique  $M(i,j)$  soit à la fois fonction :

- de la position de  $D_j$  dans le vecteur calculé,
- de la valeur de probabilité attribuée à  $D_j$ ,
- du degré d'incertitude du vecteur des probabilités.

Mais pareille étude n'a pas encore abouti.

L'analyse de cette matrice de coïncidence peut être précisée par le calcul d'autres paramètres.

### 6.2.2. La spécificité.

Elle se définit comme suit, pour chaque diagnostic.

$$\text{Sp}(D_j) = \frac{\text{VN}(D_j)}{\text{N}(D_j)}$$

où  $\text{VN}(D_j)$  (Vrais négatifs) est le nombre de malades de diagnostic autre que  $D_j$  pour lesquels l'ordinateur n'a pas porté ce même diagnostic  $D_j$ .

$\text{N}(D_j)$  représente le nombre de patients pour lesquels le diagnostic réel n'est pas  $D_j$ .

Le complément de la spécificité ( $1 - \text{Sp}(D_j)$ ) mesure, par diagnostic, le "silence" du processus, c'est-à-dire le pourcentage d'erreurs par défaut.

### 6.2.3. La sensibilité.

Elle se définit pour chaque diagnostic, elle aussi.

$$\text{Se}(D_j) = \frac{\text{VP}(D_j)}{\text{P}(D_j)}$$

où  $\text{VP}(D_j)$  (Vrais positifs) représente le nombre de malades de diagnostic réel  $D_j$  correctement calculé.

Le complément de ce paramètre mesure, quant à lui, pour chaque diagnostic  $j$ , le "bruit" du processus, c'est-à-dire le pourcentage d'erreurs par excès.

### 6.2.4. La confiance.

$$\text{Con}(D_j) = \frac{\text{VP}(D_j)}{\text{NC}(D_j)}$$



où  $NC(D_j)$  est le nombre de malades pour lesquels le diagnostic calculé est  $D_j$ .

La détermination de ces différents paramètres apporte au médecin responsable des critères de jugement susceptibles de corroborer ou de mettre en doute la connaissance de chaque maladie qu'il a dû traduire au préalable en termes de probabilités.

### 6.3. Apport pédagogique.

La réalisation et l'utilisation toute relative de ce système a tout de même permis une prise de conscience de trois notions essentielles.

Trop souvent, l'enseignement médical se fait "en colonne", c'est-à-dire diagnostic par diagnostic ou maladie par maladie, alors que la pratique médicale s'opère "en ligne": "Quelles sont les éventualités susceptibles de provoquer l'apparition de tel signe ou tel syndrome?" Ce passage colonne-ligne n'est pas toujours aisé à assimiler et à appliquer, et l'un des mérites du processus bayésien est de le réaliser et de prendre en compte tous les diagnostics inégalement probables face à un signe quelconque.

La non-apparition d'un symptôme est souvent négligée dans l'enseignement et la pratique traditionnels alors qu'elle a une valeur informative non négligeable. Nous pensons ici plus spécialement au diagnostic différentiel.

Quelle que soit la fréquence d'un signe, celui-ci n'a que peu de valeur informative si sa fréquence d'apparition est sensiblement équivalente dans toutes les affections. D'autre part, même si un signe est très spécifique d'une affection, il n'a pourtant que peu de valeur s'il y est très rare.

Cette prise de conscience constitue, pour les concepteurs du système, un résultat très appréciable et finalement, un bilan positif. Ces idées sont maintenant appliquées dans leur enseignement.

## §7. Critique.

### 7.1. Estimation\_des\_paramètres\_de\_calcul.

Malgré la prise en considération de diagnostics "opérationnels", le raisonnement n'en est pas moins basé sur des valeurs d'utilité aussi subjectives que les probabilités conditionnelles.

Il faut néanmoins admettre que le principal avantage présenté par l'exploitation de cette matrice d'utilités est une certaine cohérence dans la subjectivité, une certaine homogénéité et une rigueur de raisonnement puisque les facteurs de coût fixés "à la main" font l'objet d'un certain contrôle et restent par après invariables, donc non exposés aux "sautes d'humeur" du médecin.

A ce propos, nous pouvons faire remarquer que W.I.CARD, (B13), en ophtalmologie, emploie des utilités objectives, mesurables à partir de l'amélioration ou de la diminution d'acuité visuelle de diverses interventions.

### 7.2. Problème\_de\_compromis.

Ainsi que nous l'avons déjà souligné, quels que soient leurs éléments communs, les systèmes d'aide à l'enseignement et d'aide au diagnostic ne peuvent être simultanément optimaux dans les deux optiques. Le choix des données, la détermination des probabilités a priori, le "public" auquel on s'adresse la précision recherchée, l'analyse des résultats et surtout les exigences de temps et de risque font la différence.

### 7.3. Diagnostic\_utile.

La notion de diagnostic utile, telle qu'elle est dégagée, présente l'inconvénient d'incorporer dans le processus un premier niveau d'interprétation, en dégageant de la liste des probabilités une proposition d'action qui relève en fait de la



responsabilité du médecin. La formulation de ce diagnostic opérationnel constitue une sorte d'intrusion dans la fonction médicale qui, à la limite, peut altérer la décision du médecin en l'influençant quelque peu.

## CHAPITRE 5. INFORMATIQUE ET MEDECINE.

### BILAN ET PERSPECTIVES.

Si nous excluons les applications purement administratives, nous pouvons distinguer trois grandes catégories d'application de l'informatique à la médecine. Ce sont la surveillance automatique des grands malades, les banques de données médicales et l'aide au diagnostic proprement dite. Chacun de ces domaines présente à la fois des aspects bénéfiques et des lacunes, que nous pourrions peut-être détailler.

#### A. La surveillance automatique des grands malades.

"Dans certaines maladies graves (grandes infections, intoxications, traumatismes, affections aiguës cardiaques, respiratoires ou rénales), diverses fonctions physiologiques peuvent être altérées simultanément (respiration, circulation, métabolisme...). La surveillance de ces fonctions, la correction des troubles reconnus peut devenir extrêmement complexe. On a donc développé des systèmes de surveillance automatique. Les principaux paramètres physiologiques (rythme respiratoire, rythme cardiaque, pression artérielle...) ou biochimiques sont enregistrés de façon continue ou discontinue. Les informations recueillies sont traitées par un ordinateur qui peut ainsi éditer des bilans réguliers, déclencher des alarmes et même prévoir la survenue de complications par l'analyse multidimensionnelle des données fournies. On peut envisager certains traitements (au sens médical) automatisés : par exemple, les modifications de la concentration du sucre sanguin, décelées par un système automatique de mesure, pourraient régler l'injection d'insuline, de façon à obtenir un équilibre permanent...etc." (B44)

Pareilles applications suscitent un intérêt croissant mais nécessitent, faut-il le souligner, un degré de confiance exceptionnel, une connaissance médicale unanime et une prise du risque mesurée. Avant que d'être adopté par un service



clinique, un système de ce type doit avoir fait ses preuves pendant de longues périodes de test.

Sous cette dernière condition, les facilités et les garanties offertes à l'heure actuelle paraissent déterminantes quant à la mise en oeuvre de cette surveillance automatique. Bien des services ne pourraient désormais s'en passer, si ce n'est peut-être au prix d'une embauche de personnel qualifié assez importante.

#### B. Les banques de données médicales.

A ce niveau, il nous est apparu judicieux de distinguer les banques de données relatives aux individus de celles relatives à la connaissance médicale (maladies, produits toxiques, médicaments,...)

Les premières, parmi lesquelles figureront nos dossiers d'appel des centres anti-poisons (cfr 2<sup>e</sup> partie), au-delà des délicats problèmes de fond auxquels nous avons déjà fait allusion, offrent des perspectives et des facilités de travail très encourageantes, qu'elles soient limitées à la mention des hospitalisations ou consultations ou qu'elles s'étendent aux renseignements d'ordre plus personnel. Seules, leur mise en place et leur adoption par le service clinique posent encore quelques problèmes, soulevant parfois presque des controverses de la part de médecins plus traditionnels. De notre propre avis, ces réticences se font et se feront de moins en moins vives au fil du temps et de l'implantation de nouveaux systèmes, principalement si les garanties avancées du point de vue de la diffusion des informations relatives aux individus sont respectées.

Les secondes, au sein desquelles nous classerons notre catalogue des produits (cfr 2<sup>e</sup> partie), visent à synthétiser des connaissances générales qui couvrent des domaines trop vastes pour les mettre à disposition du praticien. Le seul travail de saisie et d'enregistrement des données constitue un sérieux obstacle à la mise en place de ces banques de données. Il nécessite l'immobilisation d'une équipe scientifique

spécialisée, et ce pendant de longs mois, avant que la banque ne soit à même d'être exploitée. Une équipe doit également se charger de la mise à jour de manière à ce que toute information qui puisse s'y trouver soit la plus pertinente possible.

Malgré la mise en oeuvre d'un matériel coûteux et d'une équipe de spécialistes importante, de tels projets ont déjà vu le jour. Les banques de données individuelles (dossiers médicaux) font l'objet d'un débat qui oppose les partisans d'une centralisation intégrale à ceux d'un cloisonnement strict entre institutions médicales, entre services cliniques quand ce n'est pas entre médecins. Ce débat est loin d'être définitivement clos.

#### C. L'aide au diagnostic.

"Les systèmes d'aide au diagnostic (diagnostic automatique) sont destinés à analyser des tracés électriques (ECG,...) ou bien des ensembles de données discontinues et disparates, d'origine clinique, biologique, radiologique ...etc. Cette dernière catégorie de système n'a été jusqu'ici développée que dans des domaines très spécialisés (cardiopathies congénitales, ictères, toxicologie clinique,...)" (B44).

A notre connaissance, aucun des systèmes de cette deuxième catégorie n'est utilisé dans la pratique médicale courante, du moins en Europe. Ces systèmes en restent souvent au stade expérimental ou, s'ils le dépassent, ne servent qu'à l'analyse a posteriori des dossiers de patients. Certains y voient d'ailleurs davantage l'aspect d'aide à l'enseignement, l'étudiant pouvant réexaminer le profil d'un patient et tenter d'y poser son diagnostic.

C'est également à ce niveau que les réticences des spécialistes sont les plus fortes. La création de systèmes d'aide au diagnostic n'est acceptée, voire tolérée, que dans la mesure où le système libère le médecin d'une partie de son travail qui soit pour lui fastidieuse et n'interfère pas trop



dans la formulation même du diagnostic et dans son interprétation. Le médecin tient, ce en quoi il a parfaitement raison, à conserver le contrôle de sa propre fonction, puisqu'il en garde de toute façon l'entière responsabilité ; il tient aussi à protéger contre toute atteinte une certaine image, un certain mythe qui, depuis toujours, a fait de lui un être à part, un peu comme surestimé. Toute intrusion lui apparaît alors comme une tentative de prise de pouvoir, comme une atteinte à sa fonction, à son prestige aussi.

A côté de cet aspect, il faut bien sûr souligner le degré d'efficacité procurée par les différents systèmes mis au point. Toutes les expérimentations de ces systèmes n'ont jamais satisfait à plus de 80%, certaines ne s'en rapprochant d'ailleurs que de loin, dans la formulation du diagnostic exact. Ce chiffre totalement insuffisant doit nous faire penser à un fait : la médecine n'est faite que d'exceptions, que de cas particuliers vis-à-vis desquels une moyenne n'a que peu de valeur significative. Or, les systèmes d'aide au diagnostic basent leur raisonnement sur des données essentiellement statistiques, essayant, sans y parvenir jamais, d'intégrer des éléments subjectifs déterminants ou considérés comme tels.

La formalisation du processus de raisonnement du médecin que nous avons adoptée vaut par sa relative similitude aux processus de raisonnement de chacun. Rien ne dit et rien ne prouve qu'elle coïncide à tout coup. Libre à vous de le croire.

Enfin, il faut s'en rendre compte, tout système d'aide au diagnostic ne peut s'envisager que dans un contexte clairement défini, dans des disciplines médicales bien structurées et bien connues. Cela limite beaucoup le champ des applications et par là même les espoirs que certains avaient fondés.

IIIème PARTIE : APPLICATION A LA TOXICOLOGIE

CLINIQUE D'URGENCE.



## INTRODUCTION.

Nous venons de faire le point sur les méthodes d'aide au diagnostic. Nous les avons exposées, illustrées et critiquées, laissant entrevoir le peu de confiance qui leur était accordée au plan opérationnel par les services médicaux concernés. Si nous prolongeons notre étude, ce n'est pas pour tomber dans le même biais, pour souffrir des mêmes critiques, mais pour tenter d'adapter certains des outils que nous avons étudiés, et d'autres aussi, à un service qui en éprouve le besoin. Cette notion de besoin, voire parfois de nécessité absolue, apparaît fondamentale dans l'élaboration et la mise en place d'un système informatique.

Notre but consiste donc à analyser un service spécialisé, en l'occurrence le centre anti-poisons de Bruxelles, et à concevoir un système d'information susceptible de répondre efficacement aux besoins qu'il éprouve. Nous étudierons ces besoins au sein d'un cadre élargi à l'information toxicologique globale, cadre dans lequel nous resituerons la place et la fonction du centre anti-poisons. Au travers des modèles d'organisation de différents centres européens et des solutions informatiques adoptées par certains d'entre eux, nous approfondirons l'étude de l'environnement qui nous est proposé à Bruxelles pour en dégager une solution concrète.

Nous terminerons par les analyses fonctionnelle et organique de cette solution.

CHAPITRE 1. LA TOXICOLOGIE. FIXATION DES LIMITES DE  
RESPONSABILITE. ROLE DU C.A.P.

Dès à présent, il nous faut essayer de fixer certaines frontières qui, nous le verrons, limitent de façon plus ou moins précise les domaines de responsabilité des différents organismes chargés d'une fonction toxicologique.

Pour ce faire, nous devons nous rabattre sur une classification des intoxications qui sert de base à cette distribution des rôles que nous voudrions préciser.

A. Les intoxications aiguës.

Elles recouvrent les intoxications accidentelles, les crimes, les tentatives de suicide, ...etc.

A l'origine, c'est ce type d'intoxication qui a justifié la création des centres anti-poisons, avant même que l'on ne soit confronté de manière consciente aux autres types. Les C.A.P. (Centres Anti-Poisons) sont donc prédestinés à cette fonction d'information et de traitement des intoxications aiguës.

Au fil de ces quinze ou vingt dernières années, ils ont conservé cette orientation, l'ont précisée au niveau de la fonction qu'ils s'attribuent et qui leur est reconnue par le public.

Des études statistiques entreprises sur les fiches d'appel de différents centres ont permis de classer, par ordre d'importance décroissante :

- les tentatives de suicide et les empoisonnements criminels,
- les accidents industriels (concernant surtout les adultes)
- les absorptions accidentelles de produits ménagers (essentiellement chez les enfants),
- la pollution de l'environnement,
- la contamination par l'alimentation.



A l'heure actuelle, tous les C.A.P. disposent d'un volume de documentation considérable et riche qui leur permet de répondre aux appels de ce type dans un laps de temps très court.

#### B. Les intoxications subaiguës.

Elles couvrent essentiellement les maladies professionnelles et celles dues à l'absorption plus ou moins régulière de médicaments, les allergies, les phénomènes de 'tératogénicité', ...etc.

Au niveau de ce que nous pouvons appeler la deuxième phase du processus d'intoxication, la fonction de toxicosurveillance incombe à divers organismes.

En ce qui concerne la pathologie industrielle, les recherches entreprises et les progrès considérables accomplis par la Médecine du Travail l'ont amené à instaurer à son profit une sorte de chasse gardée. Il est vrai que, dans les vastes zonings industriels tels qu'ils existent aujourd'hui, les risques encourus par les travailleurs et les riverains étaient devenus si démesurés qu'il fallait prendre d'innombrables précautions afin d'éviter à la fois les accidents isolés et les expositions collectives. Si l'ère chimique a déjà fait trop de victimes, elle en a aussi tiré les leçons si bien que les services de Médecine du Travail remplissent d'ores et déjà leur rôle de traitement et ceux non moins importants de protection et de prévention, dans la mesure de leurs possibilités.

Au niveau de la pathologie médicamenteuse, il s'avère que, dans certains pays, parallèlement à l'action menée par les C.A.P., se sont établis des centres de pharmacovigilance à l'initiative d'organismes publics et professionnels (industrie pharmaceutique, Santé Publique, ...).

#### C. Les intoxications chroniques.

Elles recouvrent, entre autres, les phénomènes de cancérogenèse, de mutagenèse, les variations allergologiques et

immunologiques, les imprégnations toxiques, ...etc.

Cette troisième phase est à la fois la plus difficile et la plus spécifique. Ses principales caractéristiques sont :

- la très longue période d'observation nécessaire,
- le nombre de facteurs d'influence qu'il faut cerner,
- l'existence de limites à la connaissance scientifique.

Bien souvent en effet, il s'agit de cas sans précédents.

- les multiples différences des conditions de vie et l'interférence d'habitudes anormales (alcoolisme, tabagisme, ...). L'établissement de centres de traitement et de prévention des cas d'intoxication chronique présuppose l'existence de données abondantes et fiables au niveau des cas subaigus et subchroniques, sans perdre de vue que de solides connaissances en ces domaines ne constituent encore qu'un préalable, qu'une base de travail indispensable. Il existe en fait une gradation entre les trois phases que nous venons de mentionner.

C'est aussi surtout sur ce plan, où la pauvreté de l'information disponible fait insulte à la complexité et à l'urgence des problèmes, qu'une coopération des experts de tous les centres, de toutes les spécialités toxicologiques, à tous les niveaux, se révélerait à coup sûr bénéfique.

Au vu de cette classification plus rationnelle que pratique, le rôle du C.A.P. et sa position vis-à-vis des autres centres de toxicosurveillance semble se dessiner.

Il ne fait aucun doute qu'au niveau des intoxications aiguës, les C.A.P. sont les organismes les mieux placés car, mis en place à cette fin, ils ont acquis une expérience irremplaçable, en synthétisant les véritables expérimentations dont ils étaient les témoins, de manière à en exploiter tous les éléments à des fins utiles. Leur place est faite, leur efficacité a été démontrée et ne souffre plus de discussions.

Au niveau des cas subaigus et subchroniques, l'absence d'organismes adéquats a reporté, durant bien des années et



un peu en désespoir de cause, les demandes sur les centres anti-poisons, en ce qui concerne la pathologie médicamenteuse tout au moins. De leur côté, les services de Médecine du Travail ont progressivement pris la mesure sur ces recours désordonnés.

La meilleure solution consisterait sans doute, dans un premier temps, à unir les efforts de manière à tirer au clair les connaissances acquises et, dans un deuxième temps, à spécialiser les secteurs, comme l'a fait la Médecine du Travail, en réduisant le champ au profit de la qualité, tout en laissant la porte ouverte aux indispensables échanges d'informations. Spécialiser ne veut pas dire cloisonner.

Au niveau des cas chroniques, seule l'accumulation méthodique des données en provenance des centres de dépistage, du secteur hospitalier et des autres centres de traitement, subordonnée à l'accès libre et aisé aux informations exploitées aux deux autres niveaux, peut se révéler utile et efficiente. Cela nécessite donc la collaboration de tous les médecins et autres personnes appelées à surveiller ou à traiter ces cas.

En résumé, le rôle du C.A.P. ne se limite pas mais se concentre sur les cas aigus et subaigus, pour lesquels sa tâche est d'informer et d'agir sous sa propre responsabilité. En second ressort, l'ensemble de sa documentation, de son acquis et de ses experts doit servir d'outil mis à disposition des organismes d'autres niveaux.

En troisième lieu, il doit prendre une part active à l'établissement des centres du troisième niveau, en aiguillant utilement les intoxiqués chroniques qui lui font appel ou en leur apportant une attention accrue.

La détermination d'une fonction plus précise est subordonnée aux contraintes de structure et d'organisation et à l'envergure prise par chaque centre, facteurs qui diffèrent déjà tant d'un centre à l'autre, comme nous allons le montrer.

## CHAPITRE 2. STRUCTURE DES C.A.P.

Dans ce chapitre, notre objectif est de faire un tour d'horizon de différents pays de l'ouest européen, en étudiant de manière assez brute la structure des principaux centres anti-poisons. Au travers de la description de ces centres, le lecteur pourra de lui-même se rendre compte de la difficulté inhérente à l'établissement d'un organisme européen centralisé.

### A. Pays-Bas.

- dénomination :

('Centre National Néerlandais d'information sur les poisons')

- date de création : 1963

- genèse : subdivision de l'Institut National Gouvernemental de la Santé Publique.

- structure et fonctions :

- service de documentation chargé de la constitution de fichiers, fruit du dépouillement des publications toxicologiques par des scientifiques.
- laboratoire d'analyses.
- section de pharmacologie et de toxicologie (recherches et études).

Ce centre coopère avec le Centre de Réanimation de l'hôpital universitaire d'Utrecht qui assure, quant à lui, une permanence téléphonique à l'usage exclusif des médecins et des pharmaciens.

Près de 80 % des médecins utilisateurs donnent suite à leur appel, sur base d'un questionnaire de compléments d'information qui leur est adressé par le centre. Le Centre de Réanimation comporte un service de soins intensifs qui traitait, en 1976, près de 500 intoxiqués par an.



## B. Grande-Bretagne.

- dénomination : 'National Poisons Information Service'.
- genèse : créé à l'initiative du Ministère de la Santé.
- structure et fonctions :
  - service de renseignements à l'usage des médecins pour les cas d'intoxication aiguë.
  - établissement d'un index des substances d'usage courant (composition, toxicité, symptomatologie, conduite à tenir en cas d'intoxication aiguë).

Les informations figurant dans l'index qui couvre plus de 10000 articles à l'heure actuelle s'avèrent suffisantes sur un plan pratique.

## C. Suisse.

- dénomination : 'Centre Suisse d'Information Toxicologique'.
- date de création : 1966
- fonctions :
  - permanence assurée par le centre de documentation, à l'usage des médecins.
  - mission de prévention auprès des autorités et des fabricants.
  - recueil des dossiers médicaux élaborés et envoyés au centre par les médecins le consultant.

## D. Belgique.

- dénomination : 'Centre Belge Anti-Poisons'.
- date de création : 1964
- structure et fonctions :
  - permanence téléphonique d'information toxicologique à l'usage du public et des médecins.

- laboratoires de toxicologie clinique  
(analyses et recherches).

Les fiches d'appel font l'objet d'un traitement par ordinateur depuis 1969. Ce traitement débouche uniquement sur la publication d'un certain nombre de statistiques, n'offrant donc pas aux médecins du centre l'outil de travail qu'ils pouvaient espérer.

A titre indicatif, ce centre a répondu, en 1977, à près de 20 000 appels, chiffre largement dépassé en 1978.

Nous aurons l'occasion de détailler l'étude de cet organisme dans un chapitre ultérieur de notre étude.

#### E. France.

Le territoire français est couvert par un ensemble de centres anti-poisons à vocation essentiellement régionale (Lille, Nancy, Paris, Lyon, Marseille, ...), centres qui ont, pour la plupart, vu le jour au cours de ces quinze dernières années. Trois de ces centres, ceux de Paris, Lyon et Marseille ont tout logiquement pris une certaine ampleur, répondant à près de 80 000 appels par an, dont la moitié pour le seul centre de Paris. Ce chiffre montre l'importance du nombre des intoxications et ce, parallèlement au développement de l'industrie chimique médicamenteuse, agricole et 'industrielle'.

Rassurons-nous, les 80 000 appels ne correspondent pas à autant de cas de tentatives de suicides ou intoxications graves; certains appels sont faits à titre informatif ou préventif.

Les centres de Paris et de Lyon sont intégrés à des services cliniques, l'hôpital Fernand Widal et les Hôpitaux Lyonnais, qui leur permettent de traiter les intoxiqués et qui constituent par le fait même une source d'information toxicologique extrêmement riche et fiable.

Ces deux centres, avec celui de Marseille, collaborent, depuis 1973, à un projet informatique que nous analyserons en un moment plus opportun.



Bien qu'il soit difficile d'entreprendre une étude comparative des différentes structures des centres, tant leurs fonctions et leurs objectifs diffèrent, tentons cependant d'imaginer quelle pourrait être une structure idéale sur le plan de la collecte de l'information toxicologique, de son exploitation, de sa restitution et sur le plan du traitement des intoxications.

#### F. Le centre 'idéal'.

Comme nous l'avons vu, un C.A.P. comprend en priorité un service de documentation-information toxicologique qui met à la disposition de tous ou de certains (médecins, pharmaciens) une information clinique sur la toxicité des produits. Pour cela, il entretient un personnel médical spécialisé, susceptible de fournir par téléphone, par lettre ou lors de visites au centre, une réponse rapide et pertinente. Chaque centre dispose donc d'un volume appréciable de documentation, puisée dans la littérature et constamment remaniée grâce à l'expérience apportée par chaque cas d'intoxication. C'est dans ce but que chaque appel doit être suivi de l'envoi au demandeur, d'un formulaire de compléments d'information sur l'évolution générale du patient et sur l'efficacité du traitement qui lui a été prescrit. Dans certains cas plus complexes ou plus préoccupants, le centre se chargerait lui-même de recontacter le demandeur pour obtenir des indications plus précises.

A Paris, l'expérience montre qu'un tiers seulement des demandes font l'objet de pareilles confirmations.

Mais, pour être davantage au coeur du problème, tout centre d'information devrait se trouver greffé à :

- un centre de traitement des intoxications, ce qui lui permettrait de disposer de dossiers extrêmement complets et fiables sur des cas critiques,
- un stock d'antidotes et de médicaments d'urgence, ainsi qu'un service d'acheminement,
- un laboratoire d'analyses toxicologiques,
- un service de prévention et d'épidémiologie,

- un centre de recherches scientifiques.

La collaboration étroite de ces différents secteurs doit permettre : - une meilleure connaissance des produits et des risques encourus lors d'intoxications, et des traitements appropriés, donc une action plus efficace auprès du public, car toute observation constitue une véritable expérimentation clinique qui doit d'autant plus être utilisée qu'une étude des toxiques sur l'homme ne peut évidemment pas être envisagée dans d'autres conditions.

- une centralisation idéale des sources de données et de leur stockage, indispensable à la bonne marche de chaque centre.

Mais ne nous attardons pas sur cette tentative de description d'une organisation et d'un système d'information qui n'existent pas; revenons-en à une situation concrète pour analyser les perspectives qu'offre aujourd'hui l'ère de l'informatique.

Déjà implantée dans certains centres, à l'étude dans d'autres, quelles sont ses possibilités, ses limites et ses contraintes ?



### CHAPITRE 3. L'IMPACT DE L'INFORMATIQUE AU NIVEAU DES C.A.P.

Dans ce chapitre, nous cernerons le problème posé, c'est-à-dire l'opportunité d'insérer un modèle d'information automatisée dans un centre anti-poisons. Puis, nous entreprendrons une première approche dans le choix des informations à mémoriser et, à ce sujet, nous analyserons l'option informatique prise par trois centres français en ce domaine, avant de l'extrapoler au centre de Bruxelles.

#### A. Caractéristiques générales de l'information toxicologique.

L'information toxicologique présente certaines caractéristiques qui en font un domaine apte à l'automatisation :

- volume considérable d'informations à recueillir et à tenir à jour,
- dans certains cas, urgence extrême de la réponse,
- possibilité d'une vérification des diagnostics en dehors du contexte d'urgence,
- richesse des tableaux cliniques.

Cependant, il faut dès à présent se rendre compte que l'enregistrement de l'entièreté de l'information relative à de nombreux produits entraînerait un coût de saisie et d'exploitation qu'il faudrait justifier. Or, l'informatique ne prétend pas remplacer le médecin mais lui venir en aide, uniquement dans les secteurs où il en éprouve l'utilité. Le médecin connaît de nombreux produits, leur toxicité, leurs effets et surtout la conduite à tenir. Son expérience et son savoir lui permettent de répondre de manière satisfaisante à plus de 70 % des appels. Dans les autres cas, tantôt il éprouve un doute sur la composition, tantôt il essaye d'établir un rapprochement avec un produit qu'il connaît mieux, tantôt, dans une ignorance complète, il rassemble un maximum d'éléments, de symptômes, susceptibles de lui faire imaginer l'une ou l'autre hypothèse, qu'il lui

faut ensuite tester.

Le nombre impressionnant d'appels et leur urgence, le nombre de produits actuellement commercialisés et leur caractère toxique apportent une première justification quant à l'automatisation d'un fichier très fourni des produits facilement et très rapidement accessible.

En outre, le recours aux dossiers archivés, dans lesquels se dissimulait quantité d'information, se révèle subitement envisageable, s'il est réalisé avec cette même rapidité.

#### B. Etude d'un choix informatique.

Examinons à ce propos l'option informatique prise, il y a quelques années, par le C.A.P. de Lyon, en collaboration avec ceux de Marseille et de Paris.

##### §1. Saisie et exploitation des dossiers d'appel.

La base du travail informatique mis au point à Lyon en décembre 1973 a été l'élaboration du dossier permettant une saisie rapide et peu coûteuse. Le remplissage de ces fiches par les médecins-répondeurs devait s'accompagner d'une très grande minutie et d'un contrôle rigoureux. Les centres de Lyon et de Marseille ont ainsi alimenté une base-test propre à l'expérimentation.

Depuis 1974, le système est opérationnel à Lyon, supportant les fonctions suivantes :

- contrôle des dossiers,
- mise à jour et archivage des informations,
- exploitation statistique des données,
- étude de l'évolutivité des informations.

Parmi les résultats déjà obtenus, nous pouvons citer :

- toutes les statistiques épidémiologiques entreprises sur l'activité de chaque centre pour des périodes définies (origine des demandes, délai intoxication-apparition de symptômes-appel, ...),



- les études rétrospectives portant sur des intoxications aiguës causées aussi bien par un nouveau produit que par un ancien produit,
- les différentes recherches effectuées sur des thèmes très variés, tels que la toxicité d'une nouvelle formule médicamenteuse, la symptomatologie des intoxications alimentaires ou la fiabilité des traitements.

La coopération effective entre les trois C.A.P. porte sur :

- le recueil des dossiers,
- la mise au point informatique,
- l'interrogation,
- le partage des résultats de ces interrogations,
- l'élaboration d'un dictionnaire de syntaxe plus complet.

## §2. Automatisation du C.A.P. Lyon.

En outre, depuis 1976, les Hôpitaux Lyonnais ont développé un réseau temps-réel permettant l'accès à deux groupes de fonctions : - la visualisation de fichiers :

- fichier des dossiers d'appel,
- fichier des produits toxiques,
- fichier des médicaments de la pharmacie centrale des Hôpitaux Civils de Lyon,
- état d'occupation des lits dans les différents services de ces hôpitaux.
- la sélection de dossiers selon différents modes :
  - accès par le ou les produits en cause,
  - accès par une symptomatologie,
  - accès par la classe d'âge de l'intoxiqué,
  - accès par l'évolution de l'intoxication,
  - accès par le type d'intoxication,

débouchant sur la visualisation des numéros de dossiers, de la fréquence d'apparition des produits ou sur la fréquence d'apparition d'une symptomatologie.

## CHAPITRE 4. APPLICATION AU CENTRE ANTI-POISONS DE BRUXELLES.

### A. Rappel des fonctions essentielles.

Comment concevoir un système informatique adapté aux besoins d'un centre anti-poisons spécifique, en l'occurrence celui de Bruxelles, dont voici un rappel des caractéristiques essentielles.

- C'est un service d'information-documentation apte à assurer une permanence téléphonique, dans des cas d'urgence parfois extrême, essentiellement en ce qui concerne les intoxications aiguës.

Il s'agit donc de pouvoir exploiter valablement une documentation abondante, originellement contenue dans la littérature toujours disponible au centre et méthodiquement synthétisée en fichiers de cartes, de manipulation malaisée et peu sûre. Les renseignements nécessaires existent, la plupart du temps, mais c'est leur recherche, rendue complexe par le volume à maîtriser qui vient à poser problème.

D'autre part, l'exploitation des dossiers d'appel antérieurs, rendue possible précisément par les systèmes modernes d'information automatisée, ouvre des perspectives de travail beaucoup plus enrichissantes attendues depuis bien longtemps. Le système installé à Lyon en est une démonstration concluante.

Il reste donc à formaliser, à systématiser l'ensemble du processus informationnel, depuis la saisie des dossiers (qui est d'ailleurs actuellement déjà en place à Bruxelles) et la collecte de renseignements sur les nouveaux produits, jusqu'à leur utilisation pratique à l'aide d'un terminal installé au centre. Il nous faut ici préciser que les dossiers actuellement enregistrés le sont avec un retard de deux ans et ne servent pratiquement qu'à la production d'états statistiques d'un intérêt trop général, qui ne répondent en tout cas pas aux besoins fondamentaux des méde-



cins, à savoir une réponse ponctuelle ou synthétique rapide ainsi qu'une procédure d'interrogation très simple. La concentration de données, leur classification intelligente et leur mise à disposition facile et rapide constituent également un atout évident et un outil de choix, tant pour les recherches épidémiologiques que pour les études statistiques plus générales ou plus spécifiques.

#### B. Etude des besoins.

Après avoir mis en lumière ces caractéristiques essentielles, nous pouvons déterminer les réels besoins ressentis, dans le cadre d'une restructuration globale du système d'information du centre anti-poisons.

La démarche que nous proposons consiste à étudier les éléments de réponse que notre système doit pouvoir fournir au médecin, pour renseigner le demandeur de manière satisfaisante, de façon à mettre le doigt sur les lacunes du système actuel et sur les outils que nous devons concevoir.

La classification des appels, telle que nous la présentons ci-dessous, reste tout aussi valable pour le C.A.P. de Bruxelles que pour celui de Paris, les deux centres dans lesquels nous avons mené notre enquête.

Dès le premier abord, sur l'ensemble des appels effectifs, il faut en éliminer près de 30 %. Il s'agit d'appels bénins qui n'ont pas d'objet réel ou qui sont sans rapport direct avec la fonction du centre. Il va de soi que de tels appels ne 'méritent' pas de constituer de véritables dossiers et d'immobiliser ainsi des ressources mémoires inutilement.

Dans plus de 90 % des appels restants, le ou les produits en cause sont connus par la personne qui appelle.

Bien souvent ( $\pm 70$  %), le médecin du centre connaît, de par son expérience, la nature de ce produit, sa composition et la conduite qu'il faut appliquer (trai-

tement, antidote, ...).

Dans certains cas ( $\pm 20\%$ ), il peut arriver que le médecin éprouve des difficultés quant à l'identification des constituants du produit dont le nom lui est communiqué. Il a alors recours à la bibliothèque du centre ou plus généralement aux fichiers manuels créés et mis à jour par le centre lui-même.

Cette recherche, parfois longue et fastidieuse, risque toujours de rester sans réponse.

Il arrive également ( $\pm 10\%$ ) que le produit de nom fourni par le demandeur soit totalement inconnu du médecin du centre. Il peut s'agir alors d'un produit étranger, ramené de vacances, ou, cas le plus fréquent, d'un nouveau produit non encore répertorié dans les fiches ou les ouvrages disponibles au centre.

Le seul recours possible devient donc le fabricant du produit lui-même. Mais de telles démarches prennent du temps. Que penser de la personne qui attend, à l'autre bout du fil, un renseignement qui tarde à venir, qui peut la rassurer tout à fait ou qui, par le fait même qu'il se fasse attendre, peut entraîner des conséquences parfois dramatiques.

A ce propos, nous devons introduire ici une remarque : il n'existe pas, à l'heure actuelle, de législation obligeant les fabricants à communiquer systématiquement au C.A.P. les formules des produits qu'ils commercialisent.

Dans moins de 10 % des appels restants, le demandeur ne connaît pas avec certitude le ou les produits incriminés.

Ce peut être une fausse alerte, comme l'enfant retrouvé à côté d'un tube de comprimés vide; ce peut être une forte présomption, comme les quelques flacons retrouvés sur la table de nuit à côté de la victime; ce peut être aussi l'ignorance totale, l'inconnu découvert dans une chambre d'hôtel dans un état comateux.

Dans de pareils cas, l'expérience du médecin apparaît dé-



terminante. Par un interrogatoire soigné, il doit obtenir un maximum de renseignements précis de la part de la personne qui appelle, angoissée et hésitante, peu sûre d'elle. Son rôle est, à la fois, de rassurer et de s'informer sur la victime, le produit possible et l'entourage. L'ensemble des éléments ainsi obtenus, s'il ne permet pas au médecin de bâtir des hypothèses, à contrôler, peut être confronté ou plutôt devrait pouvoir l'être à des cas antérieurs figurant dans les dossiers soigneusement enregistrés. Cette dernière perspective ne s'envisage que dans la mesure où l'ensemble des dossiers aient été stockés et puissent être exploités à la demande, à l'aide de terminaux.

#### C. Proposition d'éléments de réponse.

Depuis la création de chacun des centres anti-poisons, une première forme de réponse à ces besoins a été apportée grâce à l'établissement des fichiers de cartes. Ces fichiers volumineux contenaient l'information directement utile voire indispensable sur les spécialités médicamenteuses, les produits industriels, ...etc.

Mais depuis plusieurs années, ce support se révèle inadapté et ce, pour diverses raisons :

- manipulation malaisée (déclassements, mélanges, ...),
- accroissement extraordinaire et permanent du nombre de produits à cataloguer, entraînant des problèmes de mise à jour, de réorganisation de fichiers finalement encombrants,
- existence d'une technologie de moins en moins coûteuse, de plus en plus répandue, aisément adaptable à des problèmes analogues, cette technologie offrant une solution efficace aux problèmes qui nous préoccupent :
- maîtrise du volume,
- exigence d'un temps de réponse court,
- nécessité d'une organisation plus stable,
- facilité de manipulation.

Cette nouvelle forme de réponse, le système automatique de

gestion de l'information, fournit ainsi une solution pertinente aux problèmes cruciaux qui se posaient, rendant même envisageables des démarches qui ne l'étaient pas auparavant, à savoir l'exploitation des dossiers d'appel dans les meilleurs délais et leur utilisation à des fins de recherches et d'études statistiques et épidémiologiques plus complexes.

#### D. Vue globale du système proposé.

Le système que nous proposons de mettre en place se décompose en trois sous-systèmes quasi-indépendants, bien que le recours de l'un à l'autre doive pouvoir s'effectuer très rapidement.

##### §1. Catalogue des produits.

Ce catalogue doit répertorier l'ensemble des produits usuels, produits pharmaceutiques, industriels, agricoles, ménagers, ou, pour ainsi dire, tous les produits susceptibles d'être la cause d'intoxications.

##### §2. Gestion des dossiers d'appel.

Ce système doit prendre en charge la saisie des informations sur une fiche d'appel, leur contrôle, leur enregistrement et surtout leur exploitation. Pour être efficace, il doit permettre la recherche directe et rapide des éléments nécessaires au médecin-répondeur au moment et sous la forme qu'il désire. Ce peut être un renseignement ponctuel ou une étude synthétique. Il doit aussi permettre des études globales sur des périodes déterminées à l'instigation d'un ministère ou d'un laboratoire de recherches.

##### §3. Programme d'aide au diagnostic symptomatique.

Dans les cas moins fréquents, mais il en existe, où le nom



du produit en cause est totalement ignoré et où les seuls éléments fournis par l'anamnèse se résument en quelques signes présentés par l'intoxiqué, pareil programme vise à proposer des hypothèses de diagnostic basées sur les fréquences d'apparition de ces signes dans les différents diagnostics ou groupes de diagnostics.

Trop souvent critiquée au plan théorique, cette méthode offre néanmoins l'avantage d'entreprendre une certaine démarche de raisonnement aboutissant sur la formulation d'hypothèses que le médecin doit s'efforcer d'infirmier ou de confirmer, en orientant ses questions auprès du demandeur.

#### E. Etude détaillée du système proposé.

##### §1. Le catalogue des produits.

Notre démarche se veut, une fois encore, basée sur l'expérience des autres, principalement en ce qui concerne le type de renseignements que le catalogue doit contenir. L'exposé des impératifs inhérents à son élaboration et la détermination des données qu'il nous sera possible d'enregistrer relativement à chaque produit, déboucheront sur l'analyse logique de ces données suivant la méthode proposée par A. Clarinval (réf.).

##### 1.1. Impératifs du sous-système.

- Généralisation à tous les produits toxiques commercialisés en Belgique.
- Renseignements précis du point de vue de leur toxicité.
- Possibilité d'un accès rapide et direct à cet ensemble d'informations.
- Possibilité de poser au système des questions de complexité variable.

---

Clarinval A. 'Méthodologie de l'analyse.'

polycopie, Institut d'Informatique-FNDP-NAMUR.

- Langage d'interrogation simple et facilement assimilable par des non-spécialistes.

- Possibilité d'insertions abondantes dans le catalogue de nouveaux produits. Il semble qu'une mise à jour 'on-line' ne s'impose pas. Au contraire, le contrôle rigoureux nécessité lors de l'entrée des informations justifie la mise en place de quelques sessions de mise à jour 'off-line' pendant lesquelles une équipe responsable peut concentrer son attention sur ce travail délicat sans aucune autre préoccupation.

Les trois premiers impératifs impliquent la mémorisation d'un volume important de données et le recours à un système automatisé. Les autres impératifs nécessitent, quant à eux, le traitement en mode conversationnel et donc l'installation de terminaux dans les locaux du centre.

## 1.2. Les données.

### 1.2.1. Les produits.

Les produits qui doivent figurer dans le répertoire sont :

- les principes actifs purs,
- les produits pharmaceutiques,
- les produits industriels,
- les produits agricoles (insecticides, pesticides),
- les produits ménagers (cosmétiques, ...),

ce qui revient à dire tout produit d'usage courant dans notre pays.

Si nous prenons l'exemple des médicaments, cela concerne toutes les spécialités fabriquées et commercialisées en Belgique, en France, en Allemagne et dans les autres pays limitrophes. Pour nous fixer un ordre de grandeur, il est bon de savoir que le nombre de spécialités pharmaceutiques actuellement commercialisées en France approche le chiffre de 10 000. Le nombre de principes actifs purs, éléments constitutifs de tous ces produits, avoisine, quant à lui, le chiffre de 3 000, toujours en France.



Dans cette étude, nous n'envisagerons pas les intoxications par les plantes, les champignons, les morsures et les piqûres d'animaux, qui constituent un domaine trop spécialisé.

### 1.2.2. Les informations relatives à ces produits.

La première option que nous proposons est de profiter de l'occasion qui nous est offerte pour rassembler et stocker en mémoire toutes les données relatives à ces produits. Pareille base de données aurait ainsi la vocation d'intéresser un nombre accru de personnes.

C'est l'option qui a été choisie en France par la Banque d'Information Automatisée sur les Médicaments (système BIAM). L'objectif visé par les concepteurs de BIAM dépasse de loin celui de notre centre anti-poisons mais sa vocation un peu universelle fait qu'il est intéressant de nous attarder quelque peu, au travers de la composition des fichiers et de leur structuration, sur les principales fonctions qui y sont réalisées.

#### Le système BIAM.

##### a. base de données :

Elle est organisée en deux fichiers principaux, le fichier-substances qui comporte des renseignements très complets sur les principes actifs à l'état brut.

##### b. composition des fichiers :

Elle est présentée en annexe 2.

##### c. le système :

Le système informatique mis au point pour le traitement des informations comprend deux fonctions, l'édition de tout ou partie des caractéristiques relatives à l'un ou plusieurs médicaments ou principes actifs et l'interrogation sélective, c'est-à-dire la recherche de médicaments qui possèdent certaines caractéristiques déterminées.

A ces deux fonctions s'ajoutent celle de mise à jour,

celle d'édition de dictionnaires, ...etc.

La collecte et l'enregistrement de toutes ces données représentent bien évidemment un travail considérable.

A l'heure actuelle, toute cette information existe, en grande partie, sur des fiches manuelles. Nous pourrions envisager la possibilité d'introduire globalement tous ces renseignements de manière un peu désordonnée. Mais là n'est pas notre objectif. Ce que le médecin désire, ce n'est pas se remémorer l'intégralité de la connaissance sur un produit ou sur un groupe de produits, c'est obtenir rapidement un élément objectif sur lequel il fonde son diagnostic, plus généralement la réponse qu'il donne à l'appel.

Cet élément peut être une dose, un symptôme, une composition, parfois un mécanisme d'action ou une formule chimique.

Il nous faut donc faire la part des choses. Notre but consiste à mettre en place l'outil voulu par le médecin et cet outil traite avant tout des éléments ponctuels, plus qu'une connaissance scientifique générale.

Pour mieux nous faire comprendre, peut-être est-il opportun de prendre un exemple.

En annexe 3, nous avons présenté un exemplaire de fiche manuelle. Parmi les rubriques qui y sont présentées, seules les suivantes nous intéressent, dans l'optique qui est nôtre :

- nom du produit,
- ses synonymes et ses autres appellations,
- son groupe d'appartenance,
- ses caractéristiques physiques (en partie),
- les doses thérapeutiques, décomposables en voie d'administration, dose moyenne, dose maximale,
- les contre-indications,
- la toxicité chez l'homme, chez l'enfant,
- les symptômes présentés en cas d'intoxication,
- les références bibliographiques et les sources des données,

Considérant d'autres fiches, nous aurions en outre relevé :



- la symptomatologie aiguë, subaiguë et chronique,
- le traitement approprié ou conseillé,
- le fabricant (ou le distributeur en Belgique).

Cependant, il apparaît dès à présent que certaines de ces rubriques n'existeront pas pour tous les produits. Nous devons donc définir des structures différentes et plus souples, pour laisser place aux descriptions précises et détaillées ou, au contraire, pour économiser cette place, là où les connaissances sont moins développées.

### 1.3. L'analyse des données.

"La tâche fondamentale de l'analyse des données consiste à structurer les ensembles de données, plus précisément à 'classer' les données, ce qui, au terme de la démarche, reviendra à partitionner l'espace des supports matériels utilisés par l'ordinateur.

Reste la question de choisir les représentations elles-mêmes, qui seront manipulées (...). Tâche accessoire de l'analyse des données : définir la codification de celles-ci."

A. Clarinval (m réf.)

Les étapes de l'analyse des données s'enchaînent comme suit :

- 1.- description des données individuelles.
- 2.- description des unités d'information.
- 3.- relations entre unités d'information.
- 4.- description des articles et des fichiers logiques.
- 5.- structure d'accès.

L'analyse logique débouche ensuite sur le choix d'une organisation physique où nous envisagerons :

- 1.- définition physique des données individuelles.
- 2.- organisation physique des fichiers.

### 1.3.A. Analyse logique des données.

#### 1.3.A.1. Description des données individuelles. (cfr. tableau page suivante)

##### a. Fonction d'accès (colonne 2).

Pour chaque donnée recensée (colonne 1), nous avons identifié l'objet qu'elle décrit par l'expression :

FOR EACH VALUE OF <indicatif de l'objet>

Lorsque l'on retrouve une clause du type

fabricant-code : FOR EACH<fabricant-code>

cela signifie que fabricant-code est l'indicatif primaire identifiant un objet, en l'occurrence le fabricant.

##### b. Fonction d'effectuation.

Toutes les données recensées sont des données non-calculées.

##### c. Fonction de validité.

Il est nécessaire d'indiquer les conditions d'acceptation des données primaires (non-calculées) en vue d'en permettre le contrôle.

##### - définition physique du type, format (colonne 3).

S'il s'agit d'une donnée composée, on renvoie à la liste de ses composants.

S'il s'agit d'une donnée élémentaire, on en indique le format en désignant l'alphabet de sa codification (X = alphanumérique, 9 = numérique, A = alphabétique) et la longueur du mot code.

##### - restrictions sur le type physique (colonne 4).

Pour de nombreuses données, le type réel est plus restreint que le type physique défini par le format. Il y a alors lieu de fournir, soit la liste des valeurs possibles, soit les bornes d'un intervalle de valeurs.

Dans notre cas, certains des dictionnaires de codification étant à structure ouverte, la liste des valeurs acceptables peut s'enrichir au fil du temps.



(FOR EACH)

type physique

TYPE-NOM-PROD	PRODUIT-CODE	num(1) - cfr. codif. §2	type du nom de produit
INDIC-PRIMAIRE	PRODUIT-CODE	alphanum(35) - cfr. codif. §1	nom du produit (unique)
SYNONYME	PRODUIT-CODE	id.	synonyme du nom du prod.
N-GENERIQUE	id.	id.	nom-générique du prod.
N-GROUPE	id.	id.	nom-de-groupe du prod.
PRODUIT-CODE	id.	num(7) - cfr. codif. §1	code du nom du produit
PHYS-CARACT	id.	alphanum(texte libre) - cfr. rem	caractéristiques phys.
DOSTOX-AD	id.	=(VAL, UNIT-CODE, VOIE-ADMIN-CODE)	dose tox. pour l'adulte
VAL		num(4)	valeur (000.0)
UNIT	UNIT-CODE	alphanum(10)	unité de mesure
UNIT-CODE	id.	num(2) - cfr. codif. §3	code de l'unité de mes.
DOSTOX-ENF	PRODUIT-CODE	=(VAL, UNIT-CODE, VOIE-ADMIN-CODE)	dose tox. pour l'enfant
SYMPT	SYMPT-CODE	alphanum(60)	symptôme
SYMPT-CODE	id.	num(4) - cfr. codif. §4	code du symptôme
TRAIT	PRODUIT-CODE	alphanum(texte libre) - cfr. rem	traitement
BIBLIO	id.	alphanum(100)	référence bibliographique
CONSTIT	id.	=(VAL, UNIT-CODE, PRODUIT-CODE)	constituant du produit
NOM-FIRME	FABRICANT-CODE	alphanum(20)	nom de la firme
FABRICANT	id.	=(NOM-FIRME, ADR-FIRME)	fabricant
FABRICANT-CODE	id.	num(3) - cfr. codif. §5	code du fabricant
ADR-FIRME	id.	alphanum(50)	adresse de la firme
PRES-PUBLIC	PRODUIT-CODE	(FORME-PRES-CODE, VAL, UNIT-CODE)	présentation en vente publ.
FORME-PRES	FORME-PRES-CODE	alphanum(10)	forme de présentation
FORME-PRES-CODE	id.	num(2) - cfr. codif. §6	code de la forme de prés-
CONTRE-INDIC	CONTRE-INDIC-CODE	alphanum(30)	contre-indication
CONTRE-INDIC-CODE	id.	num(3) - cfr. codif. §7	code de contre-indication
DOSTHE-AD	PRODUIT-CODE	=(VAL, UNIT-CODE, VOIE-ADMIN-CODE)	dose thérapeutique adulte
DOSTHE-ENF	id.	id.	dose thérapeutique enfant
VOIE-ADMIN	VOIE-ADMIN-CODE	alphanum(15)	voie d'administration
VOIE-ADMIN-CODE	id.	num(2) - cfr. codif. §8	code de la voie d'admin-
CAT-PRODUIT	PRODUIT-CODE	num(1) - cfr. codif. §9	catégorie de produit
NUM-REF	PRODUIT-CODE	num(2)	n° réf. biblio. relatif
MEC-PHYSIOPA	id.	alphanum(texte libre)	mécanisme physiopatholog.



d. Signification.

La dernière colonne du tableau donne la signification de chaque donnée.

e. Remarque.

Au vu de certaines fiches manuelles où sont généralement consignées des propositions de traitement et des caractéristiques physiques brutes (aspect, solubilité, ...) non structurées, il nous a semblé préférable d'autoriser le format 'texte libre' pour plusieurs rubriques. Une longueur maximum sera néanmoins définie lors de la fixation des formats physiques des pointeurs de longueur.

f. Codification.

-/f.1/ Dictionnaire des noms de produits.

Les indicatifs sont les noms par lesquels les produits sont désignés et qui serviront donc de clé d'accès aux fichiers. La nature des données et des produits nous a amené à en définir quatre types différents.

- l'indicatif-primaire : Il désigne un produit de manière univoque. L'indicatif-primaire doit donc être un nom très précis, éventuellement très long, le premier caractère étant obligatoirement alphabétique. Le choix de l'indicatif-primaire peut répondre à une préférence individuelle ou à un souci de standardisation des appellations. Dans cet ordre d'idée, il serait souhaitable, voire opportun, de choisir la D.C.I. (Dénomination Commune Internationale).

- les indicatifs-secondaires : Ce sont d'autres appellations scientifiques ou commerciales, des traductions néerlandaise ou anglaise, également utilisées pour désigner un produit de manière univoque sans que ce ne soit précisément l'indicatif-primaire. Tout indicatif-secondaire (ou synonyme) devra donc accéder, sans détours, aux



informations relatives au produit.

- les noms-génériques : Il s'agit de dénominations non univoques, dont l'emploi est assez fréquent et qui, dans l'esprit de certaines personnes, apparaissent a priori comme univoques.

Il nous faudra donc établir les correspondances existant entre l'indicatif-primaire et les noms-génériques qui lui sont associés.

- les noms-de-groupe : C'est pour une raison pratique que le lecteur découvrira par la suite que nous avons dû définir ce quatrième type de nom de produit. Un nom-de-groupe (barbiturique pour les médicaments, détergent pour les produits domestiques, ...) devrait figurer à la fois en nom-générique et en indicatif-primaire. Le barbiturique n'est pas un produit mais un ensemble de produits caractérisé par des propriétés qu'il serait fastidieux de répéter à chaque occurrence d'un élément de cette classe.

Tout nom de produit sera donc associé à un code numérique à sept chiffres, attribué selon l'ordre d'entrée dans le catalogue.

#### -/f.2/ Les types de noms de produit.

C'est un dictionnaire à structure fermée explicitant justement le type de nom dont il s'agit.

- 1...indicatif-primaire
- 2...nom-de-groupe
- 3...nom-générique
- 4...synonyme

#### -/f.3/ Les unités de mesure, de dosage.

Voici un exemple de ce que pourrait être ce dictionnaire :

- |              |                |
|--------------|----------------|
| - 1...g      | - 6...mg/h     |
| - 2...mg     | - 7...mg/crise |
| - 3...microg | - 8...mg/jour  |
| - 4...ml     | - 9...mg/prise |
| - 5...cl     | - etc...       |

-/f.4/ Les symptômes.

La codification structurée adoptée dans le catalogue des produits se retrouvera au niveau des autres sous-systèmes .

A cette fin, nous adoptons la codification préconisée par les experts européens (cfr. Banque de données européenne des C.A.P.)

- 01...troubles hépatodigestifs
- 02...signes cutanés et généraux
- 03...signes endocriniens
- 04...signes neurologiques et neuromusculaires
- 05...signes neurosensoriels
- 06...signes respiratoires
- 07...signes cardiovasculaires

Les deux autres chiffres du code numérique déterminent le symptôme en lui-même.

- exemple pour la classe 03 :

- 0300...absence de signes endocriniens
- 0301...thyroïde
- 0302...surrénale
- 0303...hypophyse
- 0399...autres signes endocriniens

Cette dernière valeur, qui laisse la porte ouverte à d'autres symptômes endocriniens non initialement prévus au niveau des fiches d'appel, n'est évidemment pas acceptée dans le catalogue des produits.

-/f.5/ Les fabricants.

Bien que le nombre de nouveaux fabricants soit extrêmement faible, ce dictionnaire doit permettre l'adjonction de nouvelles valeurs à la suite de celles qui existent déjà.

Chaque code numérique correspond à un fabricant identifié par son nom (NOM-FIRME) et son adresse (ADR-FIRME), que nous n'avons d'ailleurs pas jugé utile de décomposer. Cette adresse sera codée sur 50 caractères alphanumériques de la manière suivante : code postal/commune/numéro et nom de rue/ numéro de téléphone

-/f.6/ Les formes de présentation.

Cette rubrique a essentiellement trait aux spécialités pharmaceutiques . Les diverses formes sont reprises dans un dic-



tionnaire à structure fermée .

- exemple :

- 1...cachet, comprimé
- 2...flacon(liquide)
- 3...pommade
- 4...poudre
- ...etc.

#### -/f.7/ Les contre-indications.

A chaque code numérique correspond une contre-indication libellée en un maximum de 30 caractères. L'attribution de ces codes se fait également selon l'ordre d'apparition de valeurs non encore prévues au dictionnaire.

#### -/f.8/ Les voies d'administration.

- exemple :

- 1...voie orale
- 2...voie anale
- 3...voie nasale
- 4...voie intraveineuse
- 5...usage externe (local)
- 6...voie sous-cutanée
- 7...etc.

#### -/f.9/ Les catégories de produit.

- 1...produit de base, principe actif pur
- 2...médicament
- 3...produit industriel
- 4...produit agricole
- 5...produit domestique, ménager.

L'élaboration du contenu de ces dictionnaires de (dé)codification doit résulter d'une collaboration entre analystes et responsables scientifiques, travail fastidieux mais indispensable que nous n'avons pu réaliser. L'enquête menée au centre anti-poisons nous a permis de déceler les besoins, de formuler une proposition, de la critiquer avec les responsables du centre puis de l'analyser. Notre étude s'y limi-

tera donc, avec tout ce que cela implique, d'une part une étude minutieuse du système global et de l'autre, une approche seulement ébauchée d'éléments accessoires.

### 1.3.A.2. Description des unités d'information.

"Une unité d'information est la définition provisoire d'un article, support nécessaire à l'existence de l'information."

Toutes les données regroupées au sein d'une unité d'information décrivent un seul et même objet.

Les unités que nous avons recensées sont :

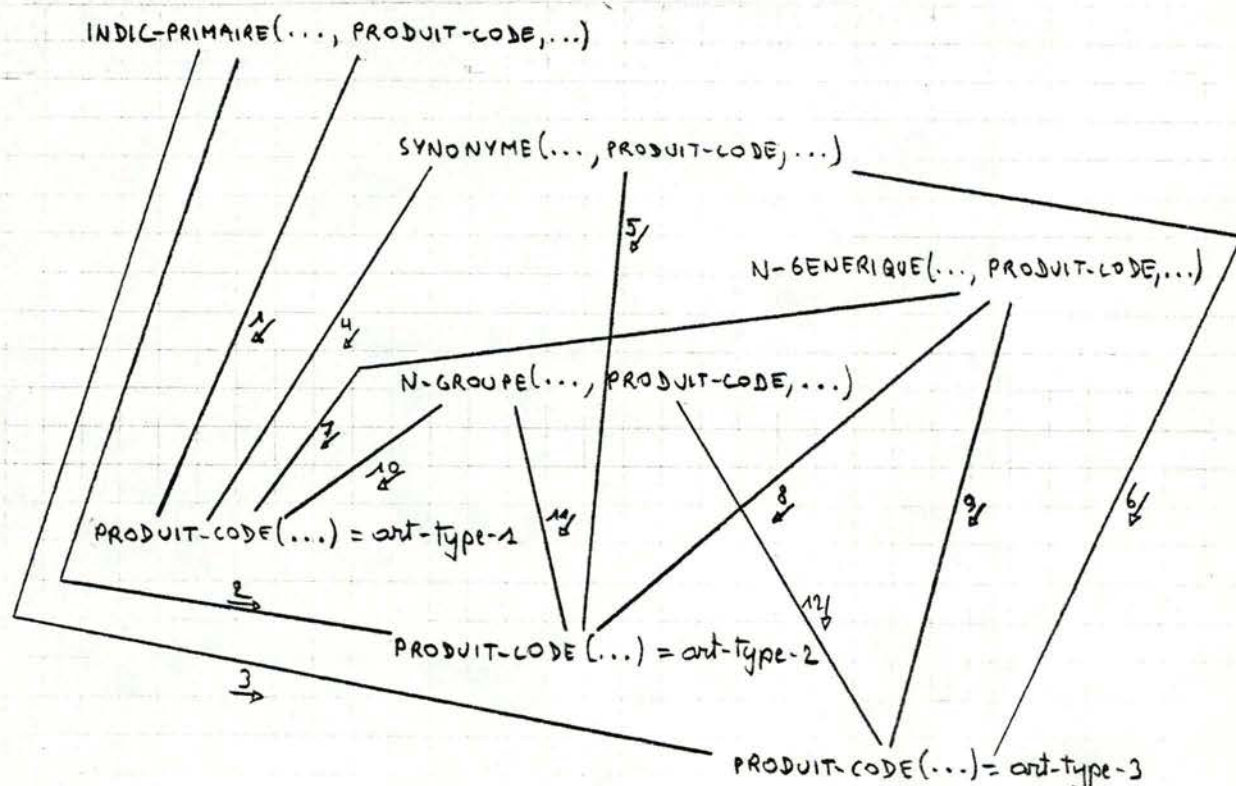
- PRODUIT-CODE(INDIC-PRIMAIRE) = art-sg-produit-1.
- PRODUIT-CODE(SYNONYME) = art-sg-produit-2.
- PRODUIT-CODE(N-GENERIQUE) = art-sg-produit-3.
- PRODUIT-CODE(N-GROUPE) = art-sg-produit-4.
- UNIT-CODE(UNIT) = art-sg-unités.
- SYMPT-CODE(SYMPT) = art-sg-symptômes.
- FABRICANT-CODE(NOM-FIRME,ADR-FIRME) = art-sg-fabricants.
- FORME-PRES-CODE(FORME-PRES) = art-sg-présentations.
- CONTRE-INDIC-CODE(CONTRE-INDIC) = art-sg-contre-indic.
- VOIE-ADMIN-CODE(VOIE-ADMIN) = art-sg-administration.
- INDIC-PRIMAIRE(PRODUIT-CODE,TYPE-NOM-PROD,CAT-PRODUIT)  
= ent-catal-type-1.
- SYNONYME(PRODUIT-CODE,TYPE-NOM-PROD,CAT-PRODUIT)  
= ent-catal-type-2.
- N-GENERIQUE(PRODUIT-CODE,TYPE-NOM-PROD,CAT-PRODUIT)  
= ent-catal-type-3.
- N-GROUPE(PRODUIT-CODE,TYPE-NOM-PROD,CAT-PRODUIT)  
= ent-catal-type-4.
- (PRODUIT-CODE,SYNONYMES,N-GENERIQUES,N-GROUPE ,TOXICITES,  
SYMPTOMATOL-TOX,BIBLIOGRAPHIE,TRAIT,PHYS-CAR,MEC-PHYSIOPA)  
= art-type-1.
- (PRODUIT-CODE,SYNONYMES,N-GENERIQUES,N-GROUPE,TOXICITES,  
SYMPTOMATOL-TOX,BIBLIOGRAPHIE,PRES-PUBLIC,THERAPEUTIQUES,  
FABRICANT-CODE,COMPOSITION,TRAIT,PHYS-CAR,EFFETS-SEC,  
MEC-PHYSIOPA) = art-type-2.
- (PRODUIT-CODE,SYNONYMES,N-GENERIQUES,N-GROUPE,TOXICITES,



SYMPTOMATOL-TOX, BIBLIOGRAPHIE, FABRICANT-CODE, COMPOSITION,  
 TRAIT, PHYS-CAR, MEC-PHYSIOPA)=art-type-3.

### 1.3.A.3. Relations entre unités d'information.

En vue de renseigner la programmation sur les voies d'accès à ménager au sein de l'organisation physique des fichiers, il importe d'établir un graphe des relations existant entre les types d'unités d'information que nous venons de recenser.



La plupart de ces relations se découvrent simplement à la redondance des indicatifs. Nous ne nous sommes pas d'ailleurs donné la peine de représenter l'ensemble des relations du type :

XXX(YYY, PPP-CODE, ...)  
 PPP-CODE(PPP)

valable pour toute donnée codée, pour nous en tenir aux relations visant les principales unités d'information.

Chacune des relations ainsi définies peut se caractériser par une propriété de cardinalité, exprimée de la façon suivante :

Soit une relation  $R$  définie entre un type  $A$  et un type  $B$  d'unités d'information. A tout moment, les contraintes suivantes doivent se trouver respectées :

- A toute unité du type  $A$  correspondent au moins  $i$  et au plus  $j$  unités du type  $B$ .
- A toute unité du type  $B$  correspondent au moins  $k$  unités et au plus  $l$  unités du type  $A$ .

La propriété se symbolise alors par un quadruplet  $(i,j;k,l)$ .

Nous pouvons appliquer cette propriété de cardinalité à l'ensemble des relations identifiées.

- relation 1 :  $(0,1;0,1)$

La flèche  $\rightarrow$  n'attribue pas un sens à la relation mais situe les types  $A$  et  $B$  à l'intérieur de cette relation.

- relation 2 :  $(0,1;0,1)$
- relation 3 :  $(0,1;0,1)$
- relation 4 :  $(0,1;0,m)$
- relation 5 :  $(0,1;0,m)$
- relation 6 :  $(0,1;0,m)$
- relation 7 :  $(0,m;0,m)$
- relation 8 :  $(0,n;0,m)$
- relation 9 :  $(0,m;0,m)$
- relation 10 :  $(0,n;0,m)$
- relation 11 :  $(0,m;0,m)$
- relation 12 :  $(0,n;0,m)$
- relation 13 :  $(1,1;0,\infty)$

#### 1.3.A.4. Description des articles et des fichiers logiques.

L'étape suivante de notre analyse consiste à répartir les types d'unités d'information que nous avons définies au sein de fichiers logiques.

"Décrire le contenu d'un fichier logique, c'est indiquer

- le(s) type(s) d'unités d'information qu'il contient,
- éventuellement les conditions de présence de ces unités,
- leur groupement, si le fichier contient des types différents,



- la fonction d'accès aux unités d'information (indicatif)."

Pour refléter la structure du fichier, nous utiliserons une suite de numéros de niveaux, le niveau 00 identifiant le fichier, le niveau 01, un type d'article de ce fichier. La structure des articles logiques est également traduite en niveaux, à la manière de la description de données du langage COBOL.

00 FICHER-DIRECT.

01 FOR EACH PRODUIT-CODE (identification de la fonction d'accès à l'article)

IF TYPE-NOM-PROD 3

IF CAT-PRODUIT = 1 = art-type-1.

02 PRODUIT-CODE pic 9(7).

02 SYNONYMES.

03 PRODUIT-CODE pic 9(7) occurs  $n_1$  times.

02 N-GENERIQUES.

03 PRODUIT-CODE pic 9(7) occurs  $n_2$  times.

02 N-GROUPE.

03 PRODUIT-CODE pic 9(7).

02 TOXICITES.

03 TOXICITE occurs  $n_3$  times.

04 VOIE-ADMIN-CODE pic 9(2).

04 DOSTOX-AD pic X(8).

04 DOSTOX-ENF pic X(8).

04 NUM-REF pic 9(2).

02 SYMPTOMATOL-TOX.

03 CAS-AIGU.

04 SYMPT-CODE pic 9(4) occurs  $n_4$  times.

03 CAS-SUBAIGU.

04 SYMPT-CODE pic 9(4) occurs  $n_5$  times.

03 CAS-CHRONIQUE.

04 SYMPT-CODE pic 9(4) occurs  $n_6$  times.

02 BIBLIOGRAPHIE.

03 BIBLIO pic X(100) occurs  $n_7$  times.

02 PHYS-CAR pic X( $n_8$ ).

02 MEC-PHYSIOPA pic X( $n_9$ ).

02 TRAIT pic X( $n_{10}$ ).

```

01  FOR EACH PRODUIT-CODE
    IF TYPE-NOM-PROD    3
    IF CAT-PRODUIT = 2      = art-type-2.
02  PRODUIT-CODE pic 9(7).
..
..      copy art-type-1 en entier.
..
02  TRAIT pic X(n10).
02  EFFETS-SEC.
    03  SYMPT-CODE pic 9(4) occurs n11 times.
02  PRES-PUBLIC pic X(8).
02  COMPOSITION.
    03  CONSTIT pic 9(14) occurs n12 times.
02  FABRICANT-CODE pic 9(3).
02  THERAPEUTIQUES.
    03  THERAPEUTIQUE occurs n13 times.
    04  VOIE-ADMIN-CODE pic 9(2).
    04  DOSTHE-AD pic X(9) occurs 2 times.
    04  DOSTHE-ENF pic X(9) occurs 2 times.
    04  NUM-REF pic 9(2).

```

```

01  FOR EACH PRODUIT-CODE
    IF TYPE-NOM-PROD    3
    IF CAT-PRODUIT    2      = art-type-3.
02  PRODUIT-CODE pic 9(7).
..
..      copy art-type-1 en entier.
..
02  TRAIT pic X(n10).
02  COMPOSITION.
    03  CONSTIT pic 9(14) occurs n12 times.
02  FABRICANT-CODE pic 9(3).

```

00 FICHER-CATALOGUE.

```

01  FOR EACH INDIC-PRIMAIRE = ent-catal-type-1.
02  PRODUIT-CODE pic 9(7).
02  TYPE-NOM-PROD pic 9(1).
02  CAT-PRODUIT pic 9(1).

```

```

01  FOR EACH SYNONYME      = ent-catal-type-2.
..  copy ent-catal-type-1.

```



```

01  FOR EACH N-GENERIQUE      = ent-catal-type-3.
    .. copy ent-catal-type-1.

01  FOR EACH N-GROUPE        = ent-catal-type-4.
    .. copy ent-catal-type-1.

00  COD-PRODUIT.
    01  FOR EACH PRODUIT-CODE  = art-sg-produit-i.

00  COD-SYMPTOME.
    01  FOR EACH SYMPT-CODE    = art-sg-symptômes.

00  COD- ...etc pour tous les autres dictionnaires de (dé)co-
      dification.

```

Remarque : Au fil de la description de la structure des trois principaux types d'article, nous avons fait référence à douze variables ( $n_i$ ,  $i=1, \dots, 12$ ), soit pour désigner le nombre d'occurrences d'un groupe répétitif, soit pour définir la longueur d'une donnée. Au niveau de l'implémentation de ces articles, nous créerons divers pointeurs où seront consignées ces différentes valeurs. Ces pointeurs une fois valorisés détermineront une structure désormais stable pour chaque occurrence d'article. Libre à l'utilisateur de déclarer un nombre d'occurrences supérieur à la réalité, dans le but de se préserver des possibilités d'insertion, au fil de l'exploitation du fichier.

#### 1.3.A.5. Structure d'accès.

Après avoir recensé et décrit les unités d'information que nous avons réparties entre fichiers logiques, après avoir identifié les relations qui existaient entre ces unités d'information, notre but est d'énumérer les accès qu'il nous faudra implémenter au stade suivant de l'analyse.

```

- ent-catal-type-1  -----  art-type-1.
  ent-catal-type-1  -----  art-type-2.
  ent-catal-type-1  -----  art-type-3.

```

Ces trois relations d'accès peuvent se résumer en une seule relation :

ent-catal-type-1 ----- art-type-i  
de cardinalité (1,1;1,1).

- ent-catal-type-2 ----- art-type-i  
de cardinalité (1,1;0,n).

- ent-catal-type-3 ----- art-type-i  
de cardinalité (1,n;0,m).

- ent-catal-type-4 ----- art-type-i  
de cardinalité (2,n;0,m). La valeur 2 peut paraître bizarre; elle s'explique cependant par le fait que tout nom-de-groupe est à la fois associé à au moins un produit du groupe (on ne déclare pas de groupe vide) et caractérisé par un ensemble de propriétés valables pour le groupe considéré dans son ensemble. Il s'agit donc d'une double relation entre des objets identiques, ce qui donnera d'ailleurs lieu à deux modes d'implémentation différents.

- Tous les accès du type

art-type-i ----- art-sg-ppp où ppp peut représenter n'importe quelle donnée codée.

- Les accès du type

ent-catal-type-j ----- art-sg-ppp.

### 1.3.B. Organisation physique des données.

L'étape finale de l'analyse doit assigner des supports aux données et ensembles de données.

#### 1.3.B.1. Définition physique des données individuelles.

Il nous faudrait procéder ici à la fixation définitive des formats des données, formats internes et externes d'entrée et de sortie. Seul le format interne fera l'objet d'une



définition plus précise, les autres formats ayant déjà été définis de manière suffisante (texte non codé).

a. Format interne. (cfr. pages suivantes)

Nous sommes parvenus à coder la majeure partie des rubriques déclarées au niveau de chaque type d'article. Les valeurs acceptées pour ces codes ont été définies dans le dictionnaire des données ainsi que leurs valeurs maximales.

Nous sommes donc en mesure de traduire en bytes (de 8 bits) la longueur de chaque type d'article.

L'abondance des groupes répétitifs et la présence de deux types de données de longueur variable nous oblige à créer, par article, une chaîne de pointeurs identifiant le nombre d'occurrences effectives de chaque donnée dans chaque occurrence d'article.

Remarques :

- Les zones fixes se trouvent en tête d'article.
- Les zones de longueur variable se trouvent en queue d'article.
- Les pointeurs se trouvent à emplacement fixe, relativement au type d'article.
- Les groupes répétitifs se situent après les zones de pointeurs.

b. Format externe d'entrée.

La solution que nous adoptons consiste à faire apparaître au vidéo une fiche d'entrée libellée (préformatée) où il suffirait de garnir les positions libres. L'entrée se fait en clair ou en code, selon l'option choisie par l'utilisateur.

Ce format doit prévoir l'extension des groupes répétitifs et la déclaration du nombre de leurs occurrences.

c. Format externe de sortie.

Il peut s'agir d'une visualisation d'une donnée ponctuelle ou d'une édition globale d'une 'fiche' de produit. Il im-

ART-TYPE-A

NB BYTES: 3

100

2

2

2

1

1

2

1

2

2

2

1

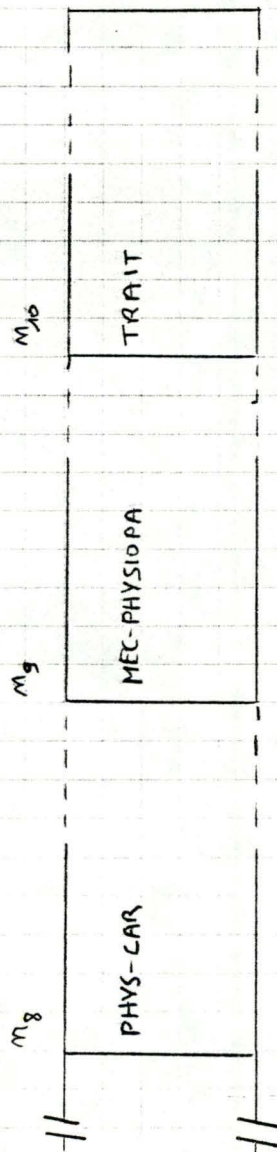
1

1

1

PRODUIT-CODE	PRODUIT-M3	M4	M5	M6	M7	M8	M9	M10	VOIE-ADMIN-CODE	VALEUR	UNIT-CODE	VALEUR	UNIT-CODE	NUM-REF	SYMPT-CODE	SYMPT-CODE	SYMPT-CODE	SYMPT-CODE	BIBLIO
ZONE DES POINTEURS										TOXICITE		CAS-AIGU CAS-SUBAIGU CAS-CHRON.							

TOXICITES =  $m_3 * \text{TOXICITE}$       SYMPTOMATOL-TOX = BIBLIOGRAPHIE =  $m_7 * \text{BIBLIO}$   
 $(m_4 + m_5 + m_6) * \text{SYMPT-CODE}$



Les pointeurs  $n_i$  peuvent éventuellement marquer l'absence de certaines rubriques ( $n_i = 0$ ).





ART-TYPE-3

NB BYTES: 3 3 1 1 1 1 1 2 2 2 2 1 2 2 1 2 1 1 2 2 2 100

PRODUIT-CODE	PRODUIT-CODE	M <sub>3</sub>	M <sub>4</sub>	M <sub>5</sub>	M <sub>6</sub>	M <sub>7</sub>	M <sub>8</sub>	M <sub>9</sub>	M <sub>10</sub>	M <sub>11</sub>	NOTE-ADNIN-CODE	VALEUR	UNIT-CODE	VALEUR	UNIT-CODE	NUM-REF					BIBLIO
		ZONE DES PONTES																			

TOXICITE

CAS-ALBU CAS-SUBSTANC CAS-CHRON.

$$\text{TOXICITES} = M_3 * \text{TOXICITE}$$

$$\text{SYMPTOMATOL-TOX} = \text{BIBLIOGRAPHIE} = M_7 * \text{BIBLIO}$$

$$(M_4 + M_5 + M_6) * \text{SYMPT-CODE}$$

6	2	M <sub>8</sub>	M <sub>9</sub>	M <sub>10</sub>
PRODUIT-CODE	FABRI-CANT-CODE	PHYS-CAR	MEC-PHYSIOPA	TRAIT
UNIT-CODE				
VALEUR				

CONSTIT

COMPOSITION = M<sub>12</sub> \* CONSTIT



porte donc de structurer la présentation (tableaux, titres) pour améliorer la lisibilité des informations décodées.

### 1.3.B.2. Organisation physique des fichiers.

La définition des articles logiques nous a amené à créer quatre types de fichiers physiques distincts.

#### a. Le fichier direct.

Il renferme les articles art-type-1, art-type-2, art-type-3, qui constituent en fait l'ensemble des renseignements stockés à propos de chaque produit. C'est donc le fichier principal garni séquentiellement, article par article, au fil des pages, autour duquel vont venir se greffer les dictionnaires de (dé)codification, le dictionnaire d'accès et le fichier inversé des noms génériques, synonymes et noms de groupe.

Nous pouvons donc représenter ce fichier direct comme une suite de pages, garnies d'articles de longueur variable et de types différents, auxquels on accèdera directement par l'adresse. Cette adresse (n° page, emplacement dans la page) connue lors de l'introduction de l'article dans le fichier direct sera mémorisée à ce moment au niveau du dictionnaire d'accès.

Il est à noter que les pages ne sont pas nécessairement contiguës. Elles sont affectées à différents fichiers suivant l'instant de leur extension physique.

PAGE 5.

ARTICLE 1	ARTI-
CLE 2	AR-
TICE 3	
	ARTICLE 4

PAGE 6.


PAGE 7.

ARTICLE 5	
	ARTICLE 6
ARTICLE 7	

PAGE 7  
DEPL. 2

POINTEURS DE PLACE LOGE.

### b. Les dictionnaires de codification.

C'est un ensemble de fichiers indexés dans lesquels s'établit la correspondance code-libellé. A tout moment, le programme peut accéder à n'importe quelle valeur d'un dictionnaire à l'aide d'un indicatif (PPP-CODE).

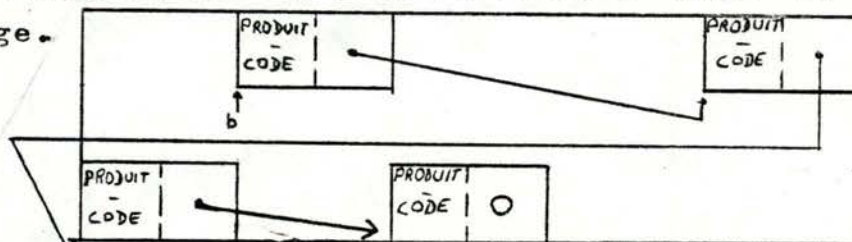
Rappelons que certains de ces dictionnaires sont à structure ouverte et que, moyennant contrôle, de nouveaux articles peuvent y être introduits.

### c. Le fichier inversé des noms-génériques et des noms-de-groupe.

Ce fichier est chargé d'établir le lien entre chaque nom-générique ou nom-de-groupe figurant au répertoire des produits et l'ensemble des indicatifs-primaires (pas de synonymes) que ce nom englobe.

Le fichier inversé est implémenté sous forme de listes simples, une liste correspondant à une association nom-générique - indicatifs-primaires.

Ce fichier occupe lui-aussi un ensemble de pages distinctes de celles des autres fichiers. L'adresse b du master de la liste, connue lors de la création d'une nouvelle liste suite à l'introduction d'un nouveau nom-générique ou nom-de-groupe dans le catalogue des produits, est également mémorisée au niveau du dictionnaire d'accès, sous la forme n° page-emplacement dans la page.



Le schéma montre que chaque maillon de la liste est constitué de deux éléments :

- un PRODUIT-CODE que le programme d'application devra traduire,
- une adresse, pointant vers le maillon suivant de cette même liste. Si cette adresse a une valeur nulle, cela signi-



fie qu'il n'existe pas de maillon suivant. L'adresse apparaîtrait toujours sous la même forme, le numéro de page n'étant pas un numéro absolu, mais un numéro de page relatif à l'ensemble des pages composant ce fichier inversé.

Nous n'avons pas fixé de limite au nombre de maillons présents dans une liste. Seul le bon sens nous dit qu'il est inutile de multiplier les noms-génériques, sous peine d'alourdir l'ensemble des performances du système. Rappelons aussi que cette notion a été introduite pour mettre le doigt sur les confusions habituelles dans les appellations des produits et donc pour les éviter.

Au niveau des noms-de-groupe, cette restriction se justifie moins.

#### d. Le fichier inversé des synonymes.

Il serait malencontreux de traiter à part les chaînes de synonymes relatifs à chaque nom de produit. Aussi définissons nous une structure analogue à celle que nous venons de décrire, à savoir une structure de listes, que nous pouvons d'ailleurs, sans aucun problème, insérer dans le même fichier physique.

L'emplacement du premier maillon de chaque liste sera, lui-aussi, mémorisé au niveau du dictionnaire d'accès.

#### e. Le dictionnaire d'accès au catalogue.

Il s'agit d'un fichier à accès par dictionnaire, constitué d'un ensemble de tables de correspondance entre valeurs de clés et adresses de pages ou d'articles du fichier direct. Les valeurs de clé sont en fait les noms de produit classés en quatre types (TYPE-NOM-PROD).

Cette distinction au niveau de la définition des types se répercutera au niveau de l'implémentation, plus exactement au niveau des pointeurs d'adresse.

Nous avons déjà créé le fichier inversé pour établir la correspondance entre nom-générique (ou nom-de-groupe) et indicatifs-primaires associés. La consultation du catalogue à

l'aide d'un nom-générique doit déboucher sur l'exploitation d'une liste du fichier inversé, alors que la consultation à l'aide d'un indicatif-primaire doit directement aboutir au fichier direct. L'introduction d'un nom-de-groupe appellera, quant à lui, l'une ou l'autre solution.

- accès par dictionnaire :

La méthode adoptée utilise, dans les tables de niveau intermédiaire, des fragments de clé de longueur fixe (8 caractères), la clé complète (35 caractères) ne figurant qu'au dernier niveau du dictionnaire d'accès.

En vue d'économiser la place mémoire et de limiter le nombre d'accès physiques, nous implémenterons une table de correspondance par page, en prenant soin d'utiliser efficacement les entrées de chaque niveau pour ménager des possibilités d'insertion d'articles au dernier niveau suffisantes.

Toute réorganisation partirait du même principe, visant à la fois à économiser le nombre total de pages occupées et la fréquence des réorganisations.

e.1. Fixation des paramètres (en système paginé).

- longueur d'une page : 2000 octets

- longueur d'un article de niveau  $n-1$ ,  $n-2$ , ..., 1 :

\* une valeur de clé (fragment) = 8 bytes

\* une adresse ( $n^{\circ}$  page) = 2 bytes

---

10 bytes

- longueur d'un article de niveau  $n$  :

\* une valeur de clé = 35 bytes

\* une adresse ( $n^{\circ}$  page, dépl.) = 4 bytes

\* TYPE-NOM-PROD, CAT-PRODUIT = 1 byte

\* une adresse = 4 bytes

\* une adresse = 4 bytes

---

48 bytes

- nombre actuel de noms de produit

à insérer dans le catalogue =  $n_1$  = + 20 000

- limite maximum raisonnable (provisoirement)

à ce nombre =  $n_2$  = 100 000

- longueur de la zone de contrôle à réserver dans



chaque page de niveau n :

* NBARTP	= 1 byte
* NBARTT	= 1 byte
	<hr/>
	2 bytes

Ces deux paramètres (NomBre effectif d'ARTicles dans la Page et NomBre d'ARTicles Triés) permettent de connaître à tout moment, pour autant que le programme connaisse la longueur d'un article de niveau n :

- la place totale occupée dans la page,
- la place libre et le nombre d'articles qu'elle peut contenir,
- le nombre d'articles triés sur lesquels peut s'effectuer une recherche dichotomique,
- le nombre d'articles non triés et leur emplacement,
- l'adresse du premier byte inoccupé.

e.2. Calcul du nombre de niveaux nécessaires.

- nombre d'articles contenus dans une page de niveau n :

$$\frac{\text{longueur utilisable d'une page}}{\text{longueur d'un article}} = \frac{1998}{48} = 41 = a$$

- nombre d'articles contenus dans une page de niveau inférieur à n :

$$\frac{1998}{10} = 199 = b$$

- nombre de niveaux à créer :

il faut choisir k tel que

$$b^{k-2} * a < n_2 < b^{k-1} * a$$

Pour k = 3, nous obtenons :

$$8\ 159 < n_2 < 1\ 623\ 641$$

e.3. Représentation du dictionnaire d'accès :

(cfr. page suivante)

2 Trois accès physiques permettent de savoir si un produit figure ou ne figure pas dans le catalogue, le

niveau 1

1 TABLE DE 199 ENTREES

clé de 8 caract.	pointeur n° page
	x
	y
	z

niveau 2

199 TABLE DE 199 ENTREES

clé de 8 caract.	pointeur n° page
	E
	v

niveau 3

39604 TABLES (POSSIBLES) DE 41 ENTREES

clé de 35 caract. (clé complète)	TYPE- NON- PROD UIT	CAT- PROD UIT	adresse1 vue F-DIRECT	adresse2 vue F-INV..	adresse3 vue F-ANNEXES
LE TYPE-NOM-PROD DETERMINE LA VALORISATION OU LA NON-VALORISATION DES TROIS ADDRESSES.					
	1		idem	very chain de synonym.	idem
	2		idem	very chain du groupe	idem
	4		/	very chain de synonym.	/
	3		/	very chain du nom- générique	/

2



quatrième permettant d'obtenir, soit les renseignements contenus dans le fichier-direct, soit ceux contenus dans le fichier-inversé, soit encore ceux contenus dans le fichier-des-annexes dont nous allons parler.

f. Le fichier des annexes.

La constitution de ce fichier répond en fait à un vœu formulé par différents médecins lors de notre enquête au centre anti-poisons. Cependant, son caractère particulier ne nous a pas permis d'en parler auparavant. Le principe sous-jacent à sa création est qu'il serait intéressant de pouvoir confier à l'ordinateur et mettre ainsi à la disposition de tous les utilisateurs du système, un ensemble de renseignements non structurables, mais d'intérêt certain. Ce sont des informations qu'il n'est pas possible de faire entrer dans la structure assez rigide de nos fichiers mais qui sont néanmoins utiles à connaître lorsque l'on est confronté à un cas d'intoxication.

L'idée qui nous est donc venue consiste à créer un fichier à structure fixe mais à contenu et longueur d'article variables. Un exemple permettra de mieux nous en faire comprendre le mécanisme.

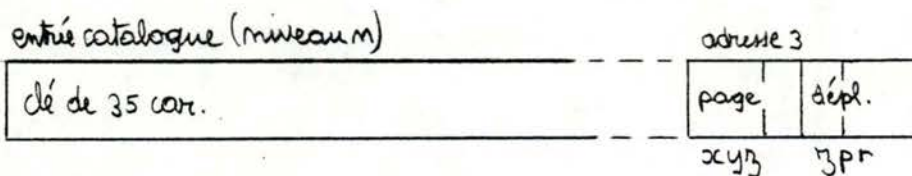
Supposons que l'un des médecins du centre découvre quelques informations intéressantes, au cours d'une lecture d'articles scientifiques. Il n'est plus ici nécessaire de souligner l'intérêt que présente la mise à disposition de ces renseignements à tous les autres médecins, au moment précis où ils en éprouveront le besoin. Le renseignement découvert est alors copié sur papier et son texte est condensé; la référence peut également y figurer. Le tout est ensuite proposé à l'équipe chargée de la mise à jour du système.

Le renseignement (référence éventuelle comprise) constitue un nouvel article du fichier des annexes, article qui sera inséré dans une structure de liste, chaque liste étant relative à un nom de produit du catalogue.

Lors de la constitution du catalogue, nous avons prévu une

adresse 3, éventuellement nulle si cette entrée n'est rattachée à aucun renseignement annexe.

Nous pouvons représenter comme suit la structure physique du fichier des annexes :



1 page de F-ANNEXES

xy3

adresse du maillon suiv.	longueur de l'art.	texte
-----------------------------	-----------------------	-------

k bytes

↳ si = 0, pas de suivant.



## 1.4. Analyse des traitements.

### 1.4.1. Création.

La création du catalogue des produits implique à la fois l'implémentation des enregistrements dans les différents fichiers et l'élaboration du dictionnaire d'accès.

Lors de cette phase de création, nous procéderons en trois étapes :

- Introduction d'un maximum d'enregistrements de façon à créer les tables de pointeurs constituant le dictionnaire d'accès au catalogue sur une base (nombre) d'articles suffisamment stable. De toute manière, une exploitation effective de ce catalogue ne peut s'envisager que dans la mesure où plus de la moitié des produits suscitant le doute y figurent.
- Création des différents niveaux du dictionnaire d'accès à l'aide de l'algorithme présenté ci-après (processus de création).
- Mise en exploitation expérimentale puis de routine, parallèlement à l'insertion de nouveaux produits.

Les phases du processus de création et l'algorithme d'élaboration des tables du dictionnaire d'accès sont inspirées de C. Cherton (réf.).

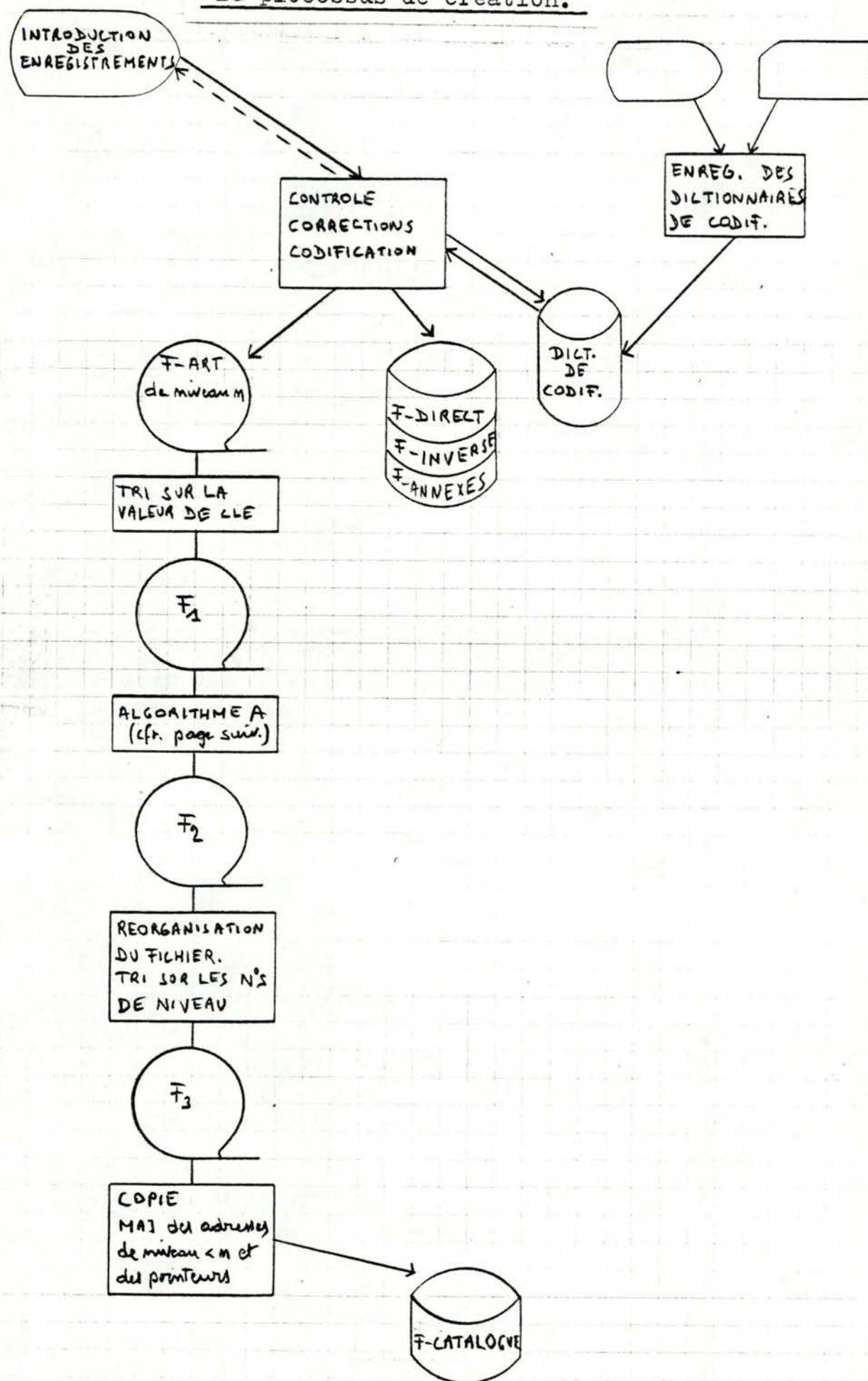
#### notations :

- $i$  : numéro de niveau ( $i = 1, \dots, n$ ).
- $n^{\circ}p$  : numéro de page.
- $p_i$  : compteur de positions dans l'enregistrement courant de niveau  $i$ .
- $n_i$  : niveau  $i$
- $f(\text{clé})$  : fragment retenu de la clé courante.

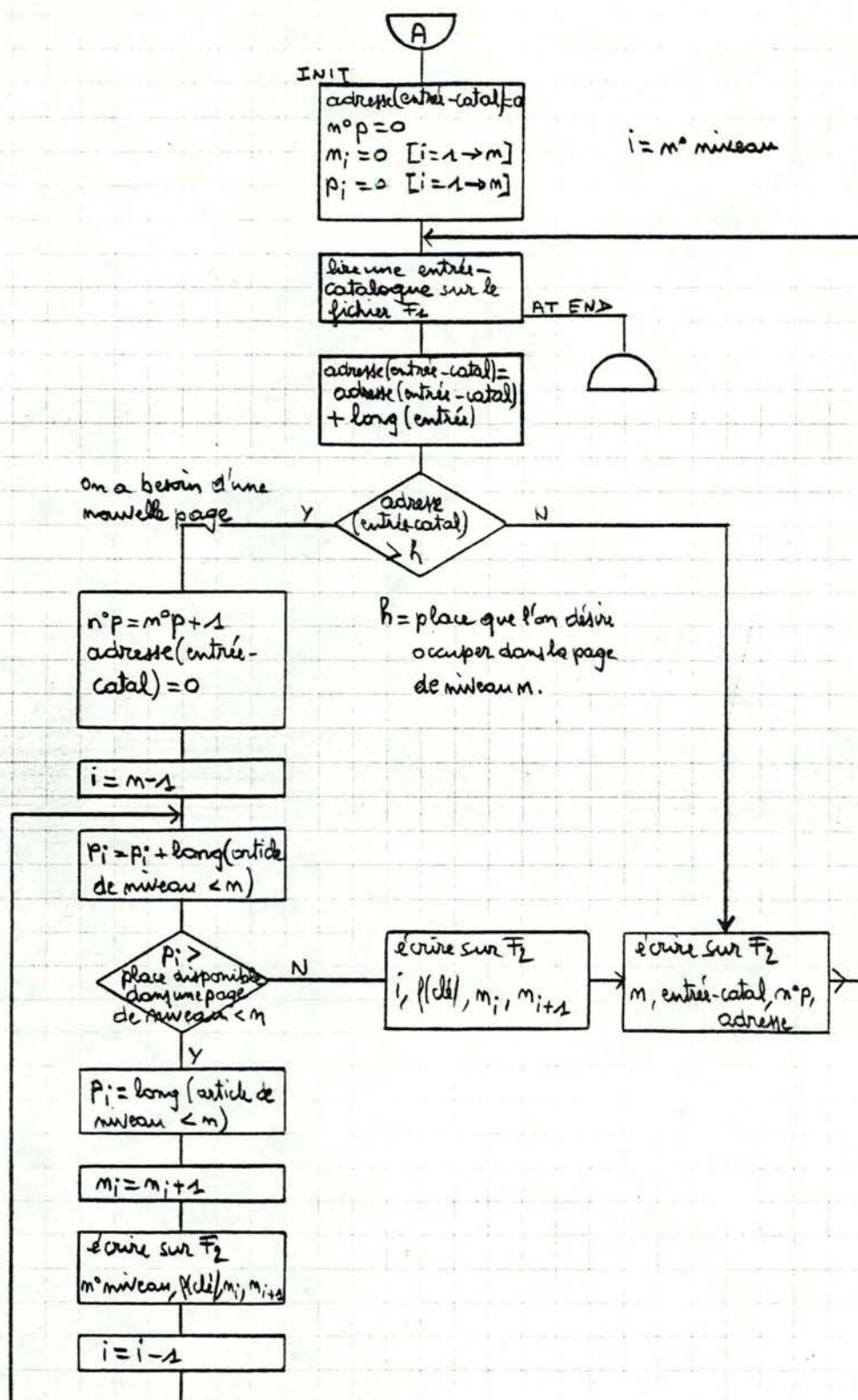
---

C. Cherton. 'Organisation des fichiers (1ère partie).'  
polycopie, Institut d'Informatique-FNDP-NAMUR (1976)

# Le processus de création.







### explication de l'algorithme :

L'algorithme A se charge de lire les enregistrements du fichier  $F_1$  (les entrées-catalogues) et de leur affecter une adresse relative (à partir de zéro) en fonction de leur longueur (fixe). Chacun de ces enregistrements se trouve recopié sur le fichier  $F_2$  accompagné d'un numéro de niveau et du numéro d'enregistrement de niveau  $n$  dans lequel il se trouve.

Grâce à  $p_i$  qui est chaque fois incrémenté de la longueur du fragment de clé retenu et de celle du pointeur correspondant, on sait si la longueur cumulée des entrées de niveau  $n$  ainsi construites reste inférieure à la taille d'un enregistrement du dictionnaire. Dès qu'il n'en est plus ainsi, on augmente  $n_i$  de 1, (indiquant ainsi que l'on passe à l'enregistrement de niveau  $n$  suivant), on réinitialise  $p_i$  (pour compter les positions occupées dans le nouvel enregistrement de niveau  $n$  que l'on commence à construire) et l'on passe au niveau précédent afin de construire l'entrée correspondante de l'enregistrement de niveau  $i-1$ . On procède de la même façon de niveau à niveau, jusqu'à ce que l'on tombe sur un enregistrement qui n'est pas encore rempli, puis on passe à l'enregistrement suivant (entrée-catalogue de  $F_1$ ).

En fin de fichier  $F_1$ , l'algorithme est terminé. On applique alors la phase suivante du processus de création.

#### 1.4.2. Mise à jour.

Chaque jour, de nouveaux produits apparaissent sur le marché, de nouveaux médicaments en remplaçant d'anciens, sous un autre nom ou selon une formule quelque peu modifiée. Comme nous l'avons vu, l'intérêt du catalogue des produits se situe dans



sa faculté de répondre rapidement à tous les problèmes qui peuvent se poser à propos d'un produit quelconque, essentiellement au plan toxicologique. L'ensemble des fichiers doit donc être tenu à jour de manière fréquente et méthodique, de façon à prévoir toute demande de renseignements au sujet de l'une ou l'autre spécialité.

De notre propre avis, une mise à jour directe n'est pas indispensable. Au contraire, il apparaît que la fonction de mise à jour devrait incomber à une petite équipe, qui consacrerait une partie de son temps au contrôle strict et à l'entrée des données sous son entière responsabilité. Il suffirait alors d'introduire en un seul lot, toutes les données dûment contrôlées.

#### a. Ajoute.

De manière classique, la création d'un nouvel article nécessite l'insertion d'une nouvelle entrée au sein du dictionnaire. Le nouvel enregistrement vient se placer en queue du fichier direct, mettant à jour le pointeur de place libre. L'entrée-catalogue est traitée comme suit :

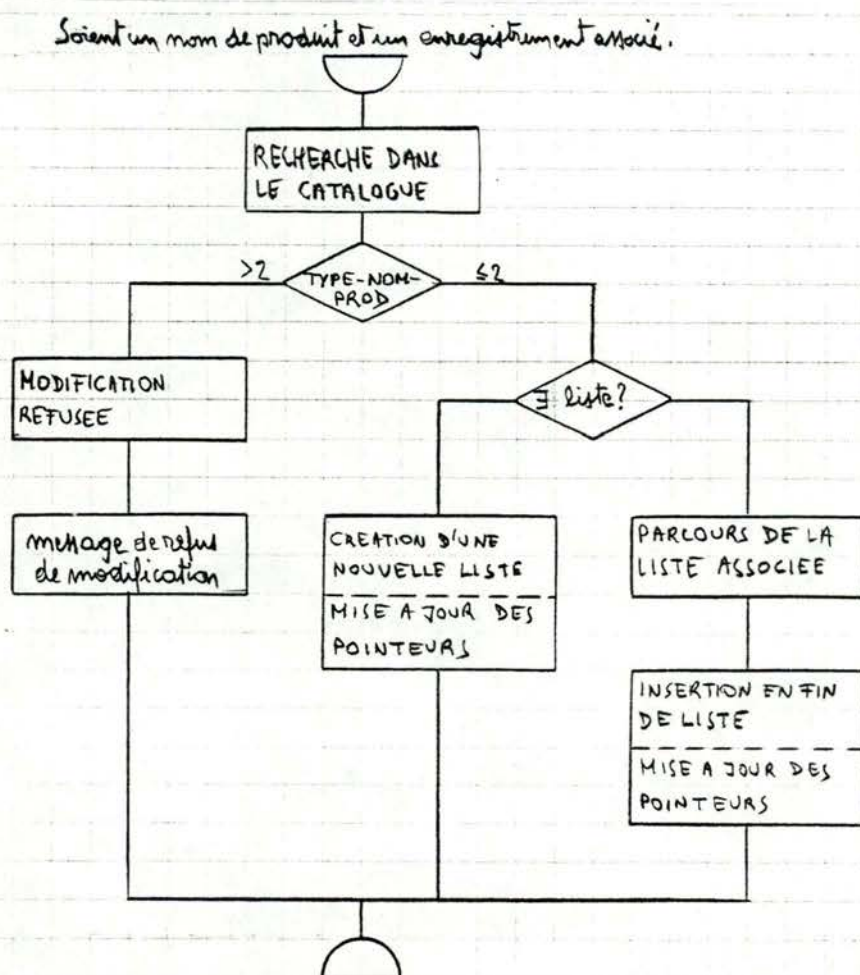
Soit la page de niveau  $n$  déterminée par la valeur de clé. La nouvelle entrée-catalogue est placée dans la page adéquate, dans la zone d'overflow interne laissée libre lors de la phase de création et de chaque phase de réorganisation. Le nombre d'articles de la page (dans la zone de contrôle de la page) est alors incrémenté. Après chaque insertion de ce type, le programme examine si la zone d'overflow interne encore libre suffit à l'insertion du prochain article de façon à lancer une éventuelle de réorganisation du dictionnaire.

#### b. Modification.

Les modifications ne peuvent porter que sur le contenu des articles; elles ne peuvent donc en supprimer, ce qui signifie qu'une modification ne peut altérer une valeur d'indicatif-primaire, ce qui supprimerait une entrée-catalogue. Pour modifier le contenu d'un article, il suffit d'y accé-

der (cfr. consultation) et de modifier les données voulues, avec l'autorisation en vigueur (mot de passe).

L'introduction d'éléments supplémentaires dans le fichier des annexes, au sein d'une liste, ou la création d'une nouvelle liste, font figure elles aussi de modification, car elles sont liées à l'entrée-catalogue d'un nom de produit. Dans un cas comme dans l'autre, il suffit d'enregistrer le nouvel élément dans les premières positions libres et de mettre à jour les valeurs des pointeurs de chaînage.



Nous prenons l'option suivante : toute liste du fichier des annexes doit être rattachée à une entrée-catalogue décrivant un INDIC-PRIMAIRE ou un N-GROUPE.

La déclaration d'un nouveau synonyme nécessite :



- la consultation sur base d'un INDIC-PRIMAIRE,
- le parcours de la liste éventuelle de ses synonymes dans le fichier inversé,
- la création et le chaînage d'un nouveau maillon, sans oublier l'introduction du nouveau nom de produit au niveau du dictionnaire d'accès.

### c. Suppression.

La suppression concerne un enregistrement entier.

La suppression d'un enregistrement peut donc entraîner celle d'un article du fichier direct, de plusieurs entrées-catalogue et de maillons du fichier inversé ou du fichier des annexes.

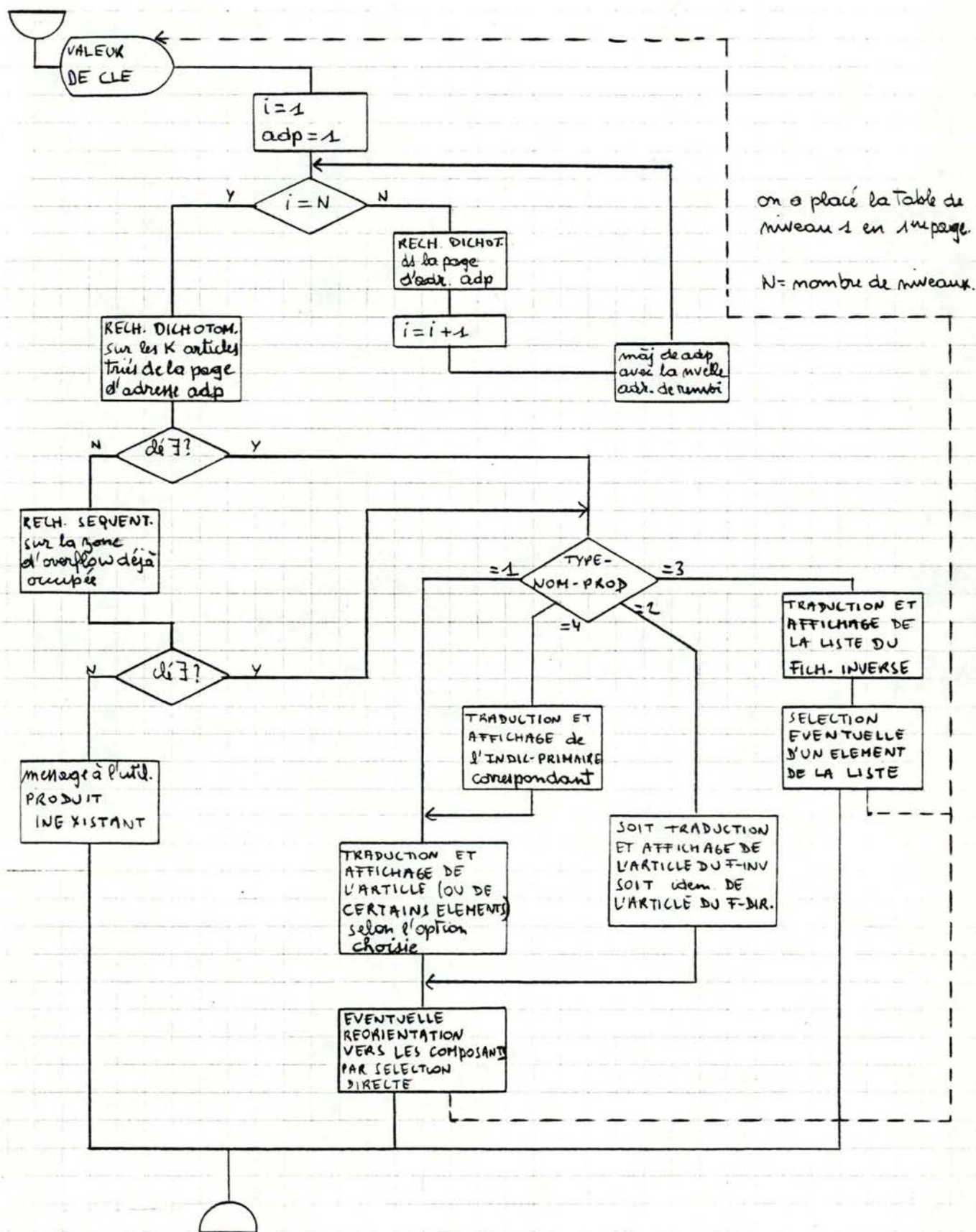
Au niveau des entrées-catalogue, la procédure ne supprime pas les entrées correspondantes mais elle rompt le lien avec l'article du fichier direct en annulant la zone adresse de l'entrée-catalogue (pointeur). L'entrée ne sera physiquement supprimée que, lors d'une phase de réorganisation, si toutefois l'on estime utile de repérer les positions libres pour économiser quelques bytes.

### 1.4.3. Consultation.

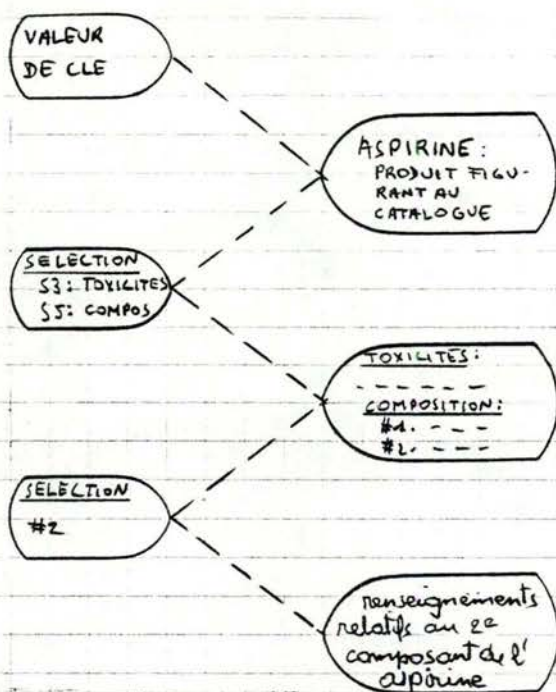
L'algorithme que nous présentons (cfr. page suivante) montre que, après sélection de la fonction de consultation, l'utilisateur introduit un nom de produit au terminal. La recherche qui s'en suit peut déboucher sur :

- le message 'PRODUIT INEXISTANT'; l'utilisateur peut alors corriger le nom introduit, constater le fait que le produit ne figure pas au catalogue ou mettre fin à sa recherche.
- l'affichage d'une liste de termes plus précis (cas du nom-générique ou du nom-de-groupe) parmi lesquels il doit se réorienter.
- l'affichage de tout ou partie des éléments demandés.

Nous présentons ci-dessous un exemple de ce que pourrait être une recherche :







Un tel système suscitant le dialogue doit faciliter toutes les démarches de réorientation des questions grâce à l'utilisation de mnémoniques, de sigles ou de codes.

#### 1.4.4. Edition.

L'édition d'une partie ou de la totalité de listes d'indicatifs, de synonymes peut être réalisée à tout moment, à partir du terminal.

Nous pouvons également éditer des parties du catalogue selon des critères variés, d'après les besoins.

#### 1.4.5. Réorganisation.

"Réorganiser un fichier est une opération qui a pour but de modifier l'implémentation physique du fichier, tout en conservant sa structure logique."

En ce qui nous concerne, le fichier direct n'est pas altéré par la réorganisation du dictionnaire d'accès, pas plus que les fichiers inversés ou le fichier des annexes.

Dans notre sous-système, le problème de la réorganisation se pose lorsqu'il s'est avéré que, dans l'une des pages du der-

nier niveau du dictionnaire, l'adjonction d'un article supplémentaire est devenue impossible. Ce fait constitue le seuil critique, détecté lors d'une phase de mise à jour.

La phase de réorganisation procède en six étapes :

- 0.- Détection du seuil critique.
- 1.- Insertion d'un maximum d'articles, puisque la phase de mise à jour n'est pas terminée, dans les pages qui le permettent encore. Certains articles sont provisoirement mis de côté.
- 2.- Tri interne sur chacune des pages du dernier niveau du dictionnaire d'accès.
- 3.- Reconstitution d'un fichier unique constitué de la suite triée des articles constituant le dernier niveau du dictionnaire. On retrouve ainsi le fichier de travail  $F_1$  de la phase de création.
- 4.- Appliquer l'algorithme de création du dictionnaire d'accès, en réajustant éventuellement les paramètres que nous avons définis ( $n_2$  et  $k$ ).
- 5.- Insérer les articles mis de côté au point 1.
- 6.- Si le nombre d'insertions effectuées au point 5 est important, on peut procéder à quelques tris internes de manière à ne pas relancer l'exploitation sur base d'un fichier déjà détérioré.

Cette phase de réorganisation n'entraîne aucune modification au niveau de l'implémentation des articles du fichier direct. L'indépendance physique est garantie par le fait que le fichier direct ne contient aucune adresse vers le dictionnaire d'accès, tout comme les autres fichiers d'ailleurs.

#### 1.4.6. Tri.

Nous n'avons pas abordé cette fonction de manière pratique. Nous ne la considérons pas comme accessoire, mais nous avons voulu nous pencher davantage sur les problèmes de structuration de fichiers que sur des problèmes purement algorithmiques. Ce ne sont d'ailleurs pas les algorithmes de tri qui manquent !



#### 1.4.7. Recherches sur le catalogue.

Dans une phase ultérieure du travail, selon les besoins ressentis, il serait utile de créer d'autres fonctions ou d'autres procédures. Citons-en quelques-unes qui nous paraissent déjà intéressantes :

- édition du catalogue de produits (en partie) à destination d'autres centres ou d'autres services,
- constitution de tables de fréquences de consultation relativement aux types de produits ou aux produits,
- sonnette d'alarme signalant un nombre accru et anormal de consultation d'un produit spécifique en un laps de temps court,
- ...

#### §2. La gestion des dossiers d'appel.

Le deuxième sous-système que nous avons identifié traite les fiches d'appel en tant que véritables dossiers médicaux. C'est en effet dans ces fiches que se concentre l'ensemble des expérimentations involontaires qui sont adressées au centre anti-poisons et c'est à partir d'elles que se profilent toutes sortes d'études tant scientifiques qu'administratives (statistiques d'activité).

Notre démarche procèdera en trois étapes. L'une, consacrée à une intéressante expérience mise sur pied par divers centres anti-poisons européens, visera à nous familiariser avec certains concepts, tout en mettant l'accent sur les buts poursuivis par cette initiative, sur l'ensemble du système en lui-même et sur les raisons qui ont provoqué son abandon.

La seconde nous resituera dans le cadre de notre centre, où nous examinerons la situation existante, en mettant le doigt sur quelques lacunes, avant de proposer, dans la troisième étape de notre travail, un système restructuré, susceptible d'aider plus efficacement le médecin dans ses tâches routinières aussi bien que dans ses recherches personnelles.

La tâche est ardue; nous nous bornerons donc à décrire sur papier le système conçu, conscients d'ignorer peut-être certains problèmes inhérents à son implémentation et à sa mise en place.

## 2.1. Les banques de données de l'INSERM.

(cfr. annexe 4)

## 2.2. Le C.A.P. de Bruxelles. Etude et critique de l'existant.

Depuis 1973, suite à l'expérience que nous venons de décrire, le centre belge, en collaboration avec le Ministère de la Santé Publique, a formalisé le recueil des dossiers d'appels téléphoniques. Ainsi donc, la tentative de centralisation des données, si elle a échoué dans ses propres objectifs, n'en a pas moins laissé de beaux restes, mis à profit par divers centres, à titre individuel.

La première étape du travail, la définition de la structure de la fiche d'appel, était déjà réalisée. Depuis lors, les années ont montré la qualité de ce travail, bien qu'inévitablement, certaines modifications aient dû y être opérées. C'est d'ailleurs dans un souci d'une perpétuelle adaptation qu'une équipe du centre revoit, chaque année, la structure des données de la fiche de façon à y faire rentrer des éléments nouveaux, non prévus ou jugés jusque là inintéressants. Jusqu'à présent, les modifications n'ont jamais altéré la structure générale du dossier, évitant par la même occasion une sérieuse révision de certains programmes d'application.

La seconde étape, le complètement des fiches lors des appels, leur contrôle par simple relecture et l'enregistrement proprement dit, s'est opérée depuis 1974, de manière plus ou moins méthodique. Cependant, l'absence de terminaux au centre et donc l'impossibilité d'une exploitation directe de ces dossiers à la demande, ont sérieusement tempéré les ardeurs de toute l'équipe. Depuis janvier 1977, les fiches de saisie s'empilent sur les bureaux des perforatrices, dans l'attente de l'octroi de subsides et de la mise à disposi-



tion d'un terminal dans les bureaux de la permanence téléphonique. Jusqu'à cette date, les médecins-répondeurs étaient astreints à remplir ces fiches, aux seules fins d'études statistiques générales d'un intérêt fort discuté.

La troisième étape, l'exploitation en tant que telle, n'a jamais vu le jour, essentiellement faute de moyens. C'est précisément la tâche à laquelle nous collaborons.

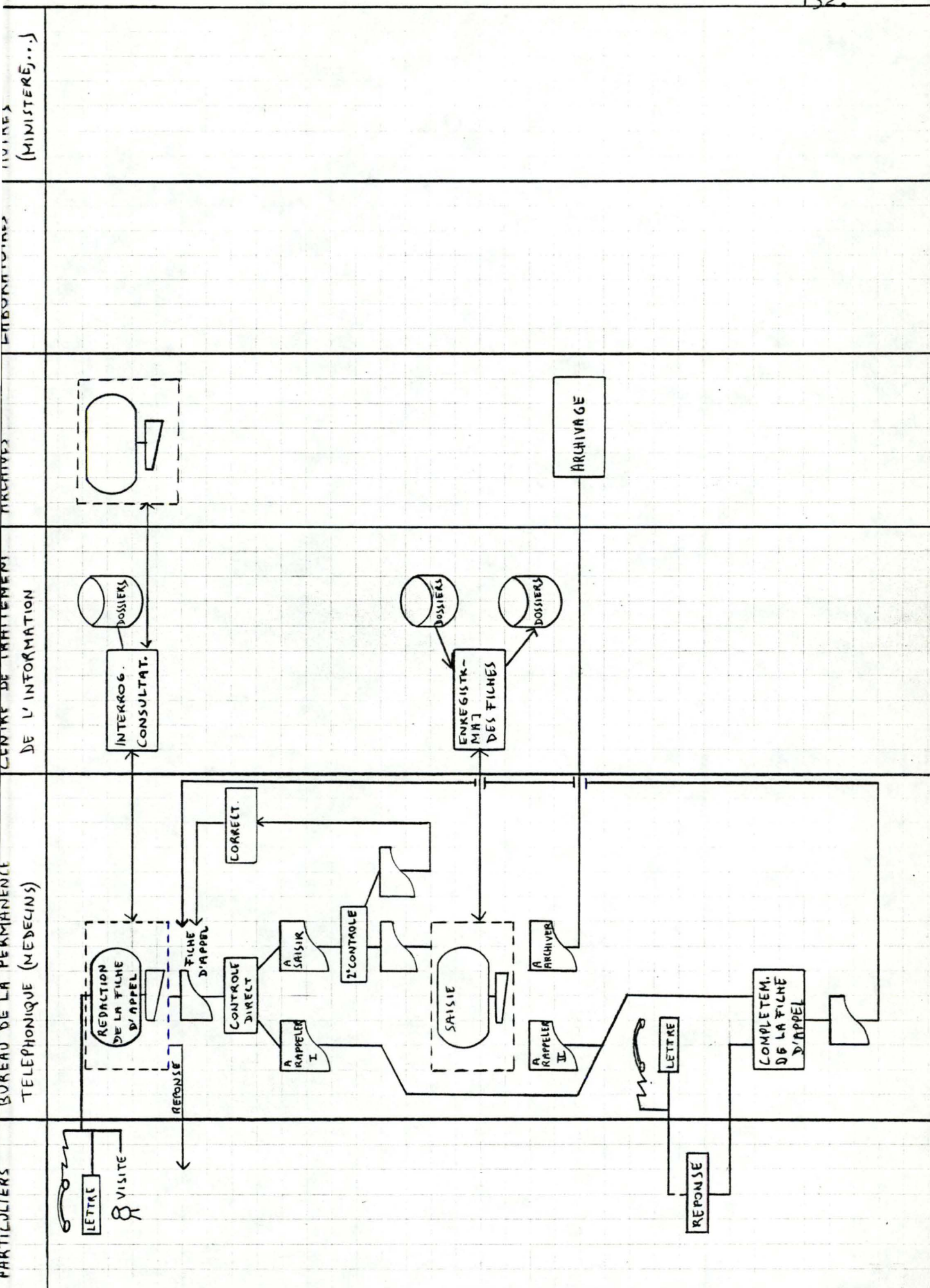
### 2.3. Présentation du nouveau système.

Nous pouvons maintenant entrer de plein pied dans la description des phases de ce sous-système. A l'image du diagramme de flux de la solution automatisée que nous proposons ci-dessous, nous aurions pu tenter d'élaborer le diagramme de flux de la solution actuelle. Cependant, l'abandon provisoire de la phase d'enregistrement des dossiers et, par conséquent, de celle d'exploitation des fichiers (statistiques), rend impossible toute formalisation des circuits d'information au sein du centre anti-poisons. Cette situation montre, de manière évidente, les vices du système actuel.

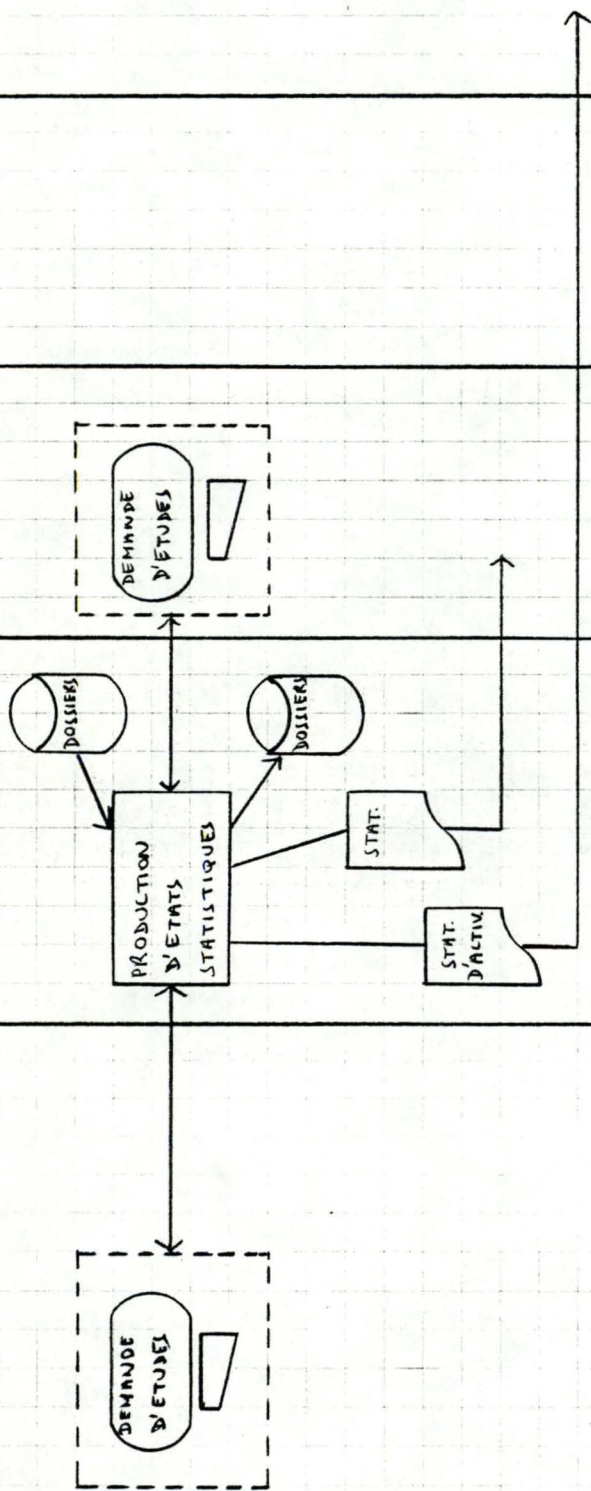
Considérons dès lors le diagramme des flux relatif au sous-système que nous concevons. Une remarque s'impose; c'est délibérément que nous n'avons représenté sur ce diagramme que les flux d'un sous-système. Nous avons voulu dissocier au maximum les sous-systèmes identifiés et montrer leur indépendance en les décrivant tour à tour.

#### 2.3.1. Constitution de la fiche d'appel.

Principalement sur appel téléphonique, mais de manière générale sur chaque demande de l'extérieur, le médecin du centre remplit une fiche d'appel. Au cours de la conversation qu'il tient avec le demandeur, le médecin peut avoir recours aux dossiers préalablement enregistrés pour argumenter sa réponse, pour la contrôler. A l'issue de l'interrogatoire, qui ne doit surtout pas en avoir la forme, la fiche d'appel se retrouve, selon les cas, plus ou moins complétée.







### 2.3.2. Contrôle direct.

Dès que le combiné est raccroché, la fiche est relue par le médecin. Il réajuste certaines réponses, transpose sur papier les éléments omis et décide de l'envoi ou du non-envoi de la fiche au poste de saisie.

Ce premier contrôle se révèle indispensable car les erreurs dues à la précipitation ou à une certaine distraction (il importe plus de répondre que de remplir la fiche) sont monnaie courante.

D'un autre côté, le choix relatif à la saisie directe relève d'un choix 'politique'. Il peut être intéressant de pouvoir se référer à des cas très récents, à propos desquels l'information est parfois incomplète. On peut également estimer fastidieux de devoir opérer, pour chaque dossier, une, voire plusieurs mises à jour là où une seule saisie et un peu de patience auraient suffi.

### 2.3.3. Deuxième contrôle.

Les responsables du centre de Bruxelles préconisent un second contrôle, opéré en fin de journée ou tous les deux jours par un expert (médecin), en vue d'assurer une cohérence globale au niveau des noms de produit (orthographe), des unités de mesure (format d'écriture des valeurs numériques), des numéros de fiche-mère (cas associés), ...etc.

Ce deuxième filtrage répond à la volonté de disposer d'une information exacte et la plus complète possible, au prix d'un minimum d'attention et de travail supplémentaire.

### 2.3.4. Saisie ou mise à jour des dossiers d'appel.

Les fiches contrôlées à saisir sont transmises au centre de saisie dont la tâche est de les encoder et de les enregistrer. Dans certains cas, l'enregistrement se résume au complètement du dossier déjà enregistré (cfr. 2.3.7.).

Les fiches d'appel sont ensuite envoyées à l'archivage, à moins qu'un indicateur ne les oblige à retourner dans les bureaux de la permanence téléphonique. D'un point de vue pra-



tique, il importe de rappeler qu'une fiche ne peut être archivée que lorsque le patient est définitivement traité, puisqu'elle fait état du bilan final.

#### 2.3.5. Archivage.

Les dossiers enregistrés et clôturés sont entreposés méthodiquement de manière à pourvoir à tout recours éventuel (rare), recours qui serait grandement facilité par l'interrogation directe des fichiers constitués.

#### 2.3.6. Consultation. Interrogation.

Il s'agit de l'aboutissement, de l'objectif qui nous est imparti. Nous aurons l'occasion de détailler les multiples possibilités de cette fonction complexe dans la suite de notre étude (analyse des traitements).

#### 2.3.7. Complètement de la fiche d'appel.

Sur lettre envoyée par le centre ou sur appel téléphonique qui leur est adressé par l'un des médecins du centre, une partie des particuliers génère une sorte de feed-back à la réponse fournie par le centre anti-poisons. Ces compléments sont intégrés à la fiche d'appel initiale, elle-même renvoyée à la phase de contrôle direct.

#### 2.3.8. Production d'états statistiques.

Déclenchée périodiquement ou sur demande expresse, cette phase produit des états statistiques à partir de l'ensemble des dossiers enregistrés. Le langage de requête doit autoriser la combinaison de tous les critères à l'aide des opérateurs logiques et arithmétiques classiques, ainsi que permettre l'édition en clair ou la représentation graphique des résultats. Il semble également qu'une certaine partie de ce travail serait facilitée par la mise en catalogue de certaines requêtes spécifiques sous forme de macros. Nous pensons ici plus spécialement aux états statistiques exigés par le Ministère de

la Santé Publique pour juger de l'activité du centre anti-poisons.

#### 2.4. Analyse logique des données.

L'annexe 5 présente un exemplaire de la fiche d'appel avec le détail de chaque type de donnée. Cette fiche, nous l'avons dit, constitue le fruit d'un travail d'experts réalisé voici près de huit ans et qui s'est avéré plus que satisfaisant, si l'on tient compte des modifications minimales qui y ont été opérées.

Nous tenterons donc de conserver les découpes existantes.

##### 2.4.1. Description des données élémentaires.

Le tableau de ces données figure à la page suivante.

##### 2.4.1.a. Codification.

###### §1. La langue.

- 0...inconnu
- 1...français
- 2...néerlandais
- 3...allemand
- 4...anglais
- 5...autre

###### §4. Le code-postal.

... que nous ne détaillerons  
pas ici !

###### §5. Le type de la demande.

- 0...type connu non anonyme
- 1...demande d'un inconnu
- 2...anonymat demandé

###### §2. Le moyen d'appel.

- 0...inconnu
- 1...téléphone
- 2...lettre
- 3...radio
- 4...télex
- 5...visite au CAP
- 6...autre

###### §6. La région.

non détaillé.

###### §3. L'urgence de l'appel.

- 0...inconnu
- 1...urgent
- 2...non urgent



d'accès		type physique	
NO-DOSSIER	id.	num(7)	numéro de dossier
NO-FICHE-MERE	id.	num(7)	numéro de fiche-mère
DATE-1	id.	=(JOUR-SEM, JJ, MM, HH, MN) - contr. 1	date-format-1
JOUR-SEM	id.	num(1) - 0 < x < 8	jour de la semaine
JJ	id.	num(2) - 0 < x < 32	jour
MM	id.	num(2) - 0 < x < 13	mois
HH	id.	num(2) - 0 < x < 24	heure
MN	id.	num(2) - 0 < x < 60	minute
LANGUE	id.	num(1) - 0 ≤ x < 6 - cfr. codif. §1	langue
CODE-MED	id.	num(2)	code du médecin-répondeur
MOY-APPEL	id.	num(1) - 0 ≤ x < 7 - cfr. codif. §2	moyen d'appel
URG-APPEL	id.	num(1) - 0 ≤ x < 3 - cfr. codif. §3	urgence de l'appel
NB-VICT	id.	num(3)	nombre de victimes
NB-VICT-DCD	id.	num(2)	nombre de victimes décédées
CODE-INT	id.	num(2)	code interne
CODE-POST	id.	alphanum(4) - cfr. codif. §4	code-postal
TYPE-DEM	id.	num(1) - 0 ≤ x < 3 - cfr. codif. §5	type de la demande
REGION	id.	num(2) - cfr. codif. §6	région
ORIG-DEM	id.	num(2) - 0 ≤ x < 18 - cfr. codif. §7	origine de la demande
* MOTIF-DEM	id.	num(2) - 0 ≤ x < 14 - cfr. codif. §8	motif de la demande
PROFESS	id.	alphanum(3) - cfr. codif. §9	profession
NATION	id.	alphanum(2) - cfr. codif. §10	nationalité
CAT-AGE	id.	num(1) - 0 ≤ x < 6 - cfr. codif. §11	catégorie d'âge
SEXE	id.	num(1) - 0 ≤ x < 3 - cfr. codif. §12	sexe
AGE	id.	=(VAL-A, UNIT-A) - contrainte 2	âge
POIDS	id.	num(3)	poids
CAT-POIDS	id.	num(1) - 0 ≤ x < 3	catégorie de poids
RACE	id.	num(1) - 0 ≤ x < 4 - cfr. codif. §14	race
VAL-A	id.	num(2)	âge (valeur)
UNIT-A	id.	num(1) - 0 < x < 4 - cfr. codif. §15	âge (unité)
GROSS	id.	=(EXIS-G, DUREE-G)	grossesse
EXIS-G	id.	num(1) - G {0,1}	existence d'une gross.
DUREE-G	id.	num(1)	durée de grossesse / inconnue



ALLERG	id.	num(1) - E {0,1}		allergie
... (CFR. fiche d'appel)	id.			...
INFECT	id.			infectieux
DATE-2	id.			date-format-2
AA	id.			année
TYPE-EXPOS	id.			type d'exposition
TYPE-EXPOS-1	id.	num(1) - E {0,1,2,3} - codif.	§16	type d'exposition-1
TYPE-EXPOS-2	id.	num(1) - E {0,1,2} - codif.	§17	type d'exposition-2
* TYPE-INTOX	id.	num(2) - E {0,...,11} - codif.	§18	type d'intoxication
LIEU-INTOX	id.	num(1) - cfr. codif.	§19	lieu d'intoxication
* APPEL-EXT	id.	num(1) - E {0,...,4} - codif.	§20	appel à l'extérieur
EX-SYMP	id.	num(1) - E {0,1,2} - codif.	§21	existence d'une symptomatol.
RESPECT-CI	id.	num(1) - E {0,1,2} - codif.	§22	respect des contre-indicat.
* TRANSPORT	id.	num(1) - E {0,1,2,3} - codif.	§23	mode de transport
* LIEU-Trait-PROP	id.	num(1) - E {0,...,5} - codif.	§24	lieu de traitement proposé
* LIEU-Trait-REEL	id.	num(1) - id.		lieu de traitement réel
* Traitement-AVT-APPEL	id.	num(2) - E {0,...,9} - codif.	§25	traitement avant appel
* Traitement-APS-APPEL	id.			traitement après appel
* Traitement-CONS	id.			traitement conseillé
Traitement	id.			traitement
PRODUIT-CODE	id.			code du nom de produit
CONN-DOSE	id.	num(7) - codif.	§27	indice de connaissance-dose
CONN-COMPOS	id.	num(1) - E {0,1} - codif.	§26	indice conn.-composition
QUANT	id.	num(1) - id.		quantité
QUANT-O	id.	= (QUANT-O, QUANT-1, ..., QUANT-5)		quantité-0
...		= (VAL, UNIT-CODE)		...
QUANT-5	id.			quantité-5
VAL	id.			valeur
UNIT-CODE	id.	num(4) - cfr. codif.	§28	code de l'unité
VOIE-ADMIN-CODE	id.	num(2) - cfr. codif.	§28	voie d'administration(code)
CAT-PRODUIT	id.	num(2) - cfr. codif.	§29	
DEM-NOUV	id.	num(1) - E {0,1,2} - codif.	§30	demande de nouvelles
REC-NOUV	id.	num(1) - E {0,1} - codif.	§31	réception de nouvelles
ANALYSE	id.	= (ANAL-LIEU, ANAL-NO)		analyse



ANAL-LIEU	NO-DOSSIER	num(1)-	c {0, ..., 7}	- codif.	§32	lieu d'analyse
ANAL-NO	id.	num(10)				numéro d'analyse
ENVOI-MED	id.	num(1)-	c {0, 1}	- codif.	§33	envoi de médicaments
RELATION	id.	num(1)-	c {0, ..., 5}	- codif.	§34	relation produit-intoxication
VALIDITE	id.	num(1)-	cfr. codif.	§35		validité du cas
EVOLUTION	id.	num(1)-	c {0, ..., 3}	- codif.	§36	évolution du cas
INTERET	id.	num(1)-	c {0, 1}	- codif.	§37	intérêt du cas
FICH-COMPL	id.	id.	id.	- codif.	§38	fiche complémentaire remplie
* SYMPT-CODE	id.	num(4)-	cfr. codif.	§39		code de symptôme
EEG	id.	num(1)-	c {0, ..., 4}	- codif.	§41	EEG
ECG	id.	num(1)-	c {0, ..., 5}	- codif.	§42	EEG
RECH-GROSS	id.	num(1)-	c {0, 1, 2}	- codif.	§40	recherche de grossesse
RADIO-CRANE	id.	id.				radio du crâne
RADIO-THORAX	id.	id.				radio du thorax
RADIO-ABDOM	id.	id.				radio de l'abdomen
RADIO-MEMB	id.	id.				radio des membres et rachis
RADIO-PREP-1	id.	id.				radio avec préparation : transit digestif
RADIO-PREP-2	id.	id.				radio avec préparation : autre
RADIOGRAPHIE	id.	id.				radiographie, radioscopie
OESOPHAGOSCOPIE	id.	num(1)-	c {0, 1, 2}	- codif.	§40	oesophagoscopie
BRONCHOSCOPIE	id.	id.				bronchoscopie
LAPAROSCOPIE	id.	id.				laparoscopie
GASTROSCOPIE	id.	id.				gastroskopie
LARYNGOSCOPIE	id.	id.				laryngoscopie
CYTOSCOPIE	id.	id.				cytoscopie
ENDO-OR-NEZ	id.	id.				exam. endoscopique oreilles-nez
ENDOSCOPIE	id.	id.				endoscopie
HISTOLOGIE	id.	id.				biopsie, histologie
NOM-TISSU-CODE	id.	num(4)-	cfr. codif.	§48		nom du tissu étudié
RES-TISSU	id.	num(1)-	c {0, 1, 2}	- codif.	§40	résultat de l'examen du tissu
EXPLO-F-RESP	id.	id.				exploration fonct. respiratoire
FOND-D-OEIL	id.	id.				fond d'oeil
* EXAM-TOXICO	id.	num(2)-	c {0, ..., 10}	- codif.	§43	matériel d'examen toxicolog.
* BIO-REIN	id.	id.	c {0, ..., 12}	- codif.	§44	examen biologique : rein
* BIO-MIL-INT	id.	id.	c {0, ..., 7}	- codif.	§45	exam. biol. : milieu intérieur

* BIO-SANG	NO-DOSSIER	num(2)-	e {0, ..., 9}	- codif. §40	examen biologique : sang
BIO-HORM	id.	num(1)-	e {0, 1, 2}	- id.	ex. biol. : explo. hormonale
VS	id.	id.			VS
UREE-SANG	id.	id.			urée sanguine
BILIR	id.	id.			bilirubine
AMYL	id.	id.			amylases
TRANSAMIN	id.	id.			transaminases
CHOLINEST	id.	id.			cholinestérases
BSP	id.	id.			Brom Sulphone Phtaléine
SPERM	id.	id.			spermogramme
AUTRE-EX	id.	id.			autres examens
AUTRES-EXAM	id.		=(VS, UREE-SANG, ..., AUTRE-EX)		antidotes-chélateurs
ANTID-CHEL	id.		=(ATROPINE, BAL, ..., AUTRE-A-CH)		atropine
* ATROPINE	id.	num(1)-	e {0, ..., 3}	- codif. §47	B.A.L.
* BAL	id.	id.			
* ... (cfr. fiche complémentaire)	id.	id.			
* ...	id.	id.			alcool éthylique
* ALC-ETYL	id.	id.			autre antidote-chélateur
* AUTRE-A-CH	id.	id.			élimination-épuration
ELIM-EPUR	id.	id.			digestive
* DIGEST	id.	num(1)-	e {0, ..., 3}	- codif. §47	
* ... (cfr. f. compl.)	id.	id.			
* ...	id.	id.			respiratoire
* RESPIR	id.	id.			neutralisation et absorption
NEUTR-ABSORP	id.		=(REDUCT, OXYD, ..., AUTRE-N-ABS)		réducteurs
* REDUCT	id.	num(1)-	e {0, ..., 3}	- codif. §47	
* ... (Cfr. f. compl.)	id.	id.			
* ...	id.	id.			autre traitement par neut-abs
* AUTRE-N-ABS	id.	id.			traitement symptomatique
TR-SYMP	id.		=(ANALEPT, ..., AUTRE-T-SYMP)		analeptiques
* ANALEPT	id.	num(1)-	e {0, ..., 3}	- codif. §47	
* ... (cfr. f. compl.)	id.	id.			
* ...	id.	id.			autre traitement symptomat.
* AUTRE-T-SYMP	id.	id.			autre traitement non classable
* AUTRE-TRAIT	id.	id.			



§7. L'origine de la demande.

- 0...inconnu
- 1...famille
- 2...public, voisins
- 3...
- ...cfr. fiche d'appel
- 16...CAP
- 17...autre

§8. Le motif de la demande.

- 0...inconnu
- 1...classique
- ...cfr. fiche d'appel
- 13...épidémiologie rens. centre

§9. La profession. (n.c.)§10. La nationalité. (n.c.)§11. La catégorie d'âge.

- 0...inconnu
- 1...adulte
- 2...enfant
- 3...prématuré
- 4...foetus
- 5...animal

§12. Le sexe.

- 0...inconnu
- 1...masculin
- 2...féminin

§13. La catégorie de poids.

- 0...inconnu
- 1...obèse
- 2...maigre

§14. La race.

- 0...inconnue

- 1...blanche
- 2...jaune
- 3...noire

§ 15. L'unité d'âge.

- 1...année
- 2...mois
- 3...jour

§ 16. Le type d'exposition-1.

- 0...inconnu
- 1...unique
- 2...à répétition
- 3...chronique

§ 17. Le type d'exposition-2.

- 0...inconnu
- 1...brève
- 2...prolongée

§ 18. Le type d'intoxication.

- 0...inconnu
- ...cfr. fiche d'appel
- 11...intoxication indéterminée

§ 19. Le lieu d'intoxication.

- 0...inconnu
- ...cfr. fiche d'appel
- 9...autre

§ 20. L'appel à l'extérieur.

- 0...inexistant ou inconnu
- 1...au producteur
- 2...à un autre CAP
- 3...autre
- 4...intervention d'un autre organisme

§ 21. L'existence d'une symptomatologie.

- 0...inconnu
- 1...nulle
- 2...présente



§22. Le respect des CI.

0...inconnu 1...avant appel 2...après conseil CAP

§23. Le transport.

0...inconnu

1...voiture ou ambulance

2...avec surveillance médicale spécialisée

3...avion ou hélicoptère

§24. Le lieu de traitement.

0...inconnu

...cfr. fiche d'appel

5...autre

§25. Le traitement(résumé).

0...inconnu

1...évacuateur

...cfr. fiche d'appel

9...mesures préventives

§26. La connaissance de la dose.

0...non 1...oui

§27. Le produit.

cfr. catalogue des produits

## §28. et §29. idem

§30. La demande de nouvelles.

0...non 1...par téléphone 2...par lettre

## §31. §33. §37. §38. 0...non 1...oui

§32. Le lieu de l'analyse.

0...non conseillée et non faite

1...faite à A.S.B.L.

...cfr. fiche d'appel

7...autre

§34. La relation.

0...inconnue    2...probable    4...nulle  
 1...certaine    3...douteuse    5...non déterminable

§35. La validité du cas.

0...inconnue    2...satisfaisante  
 1...bonne    3...nulle

§36. L'évolution.

0...inconnue    2...séquelles  
 1...guérison    3...décès

§39. Les symptômes. (cfr. catalogue des produits)

§40. 0...inconnu    1...positif    2...négatif

§41. EEG.

0...inconnu    2...plat    4...autre perturbation  
 1...normal    3...signes de localisation

§42. ECG.

0...inconnu    3...tr. de l'excitation  
 1...normal    4...tr. de conduction  
 2...tr. de repolarisation    5...cœur pulm. aigu

§43. Examen toxicologique.

0...inconnu  
 1...liquide de lavage gastrique  
 ...cfr. fiche d'appel  
 10...autre

§44. Rein.

0...inconnu  
 1...normal  
 ...cfr. fiche d'appel  
 12...autre

§45. Milieu intérieur.

0...inconnu



...cfr. fiche d'appel  
7...autre

§46. Sang.

0...inconnu  
...cfr. fiche d'appel  
9...autre

§47. Présence d'un traitement.

0...inconnu      2...conseillé  
1...avant appel   3...effectué

§48. Nom du tissu étudié.

dictionnaire à structure ouverte.

2.4.1.b. Contraintes d'intégrité.

- contrainte 1 : DATE-1 et DATE-2

JJ  $\leq$  31 si MM  $\in$  {01,03,05,07,08,10,12}

JJ  $\leq$  30 si MM  $\in$  {04,06,09,11}

JJ  $\leq$  29 si MM  $\in$  {02}

- contrainte 2 :

si CAT-AGE = 1 , alors UNIT-A = 1

si CAT-AGE = 3 , alors UNIT-A = 3

#### 2.4.2. Description des unités d'information.

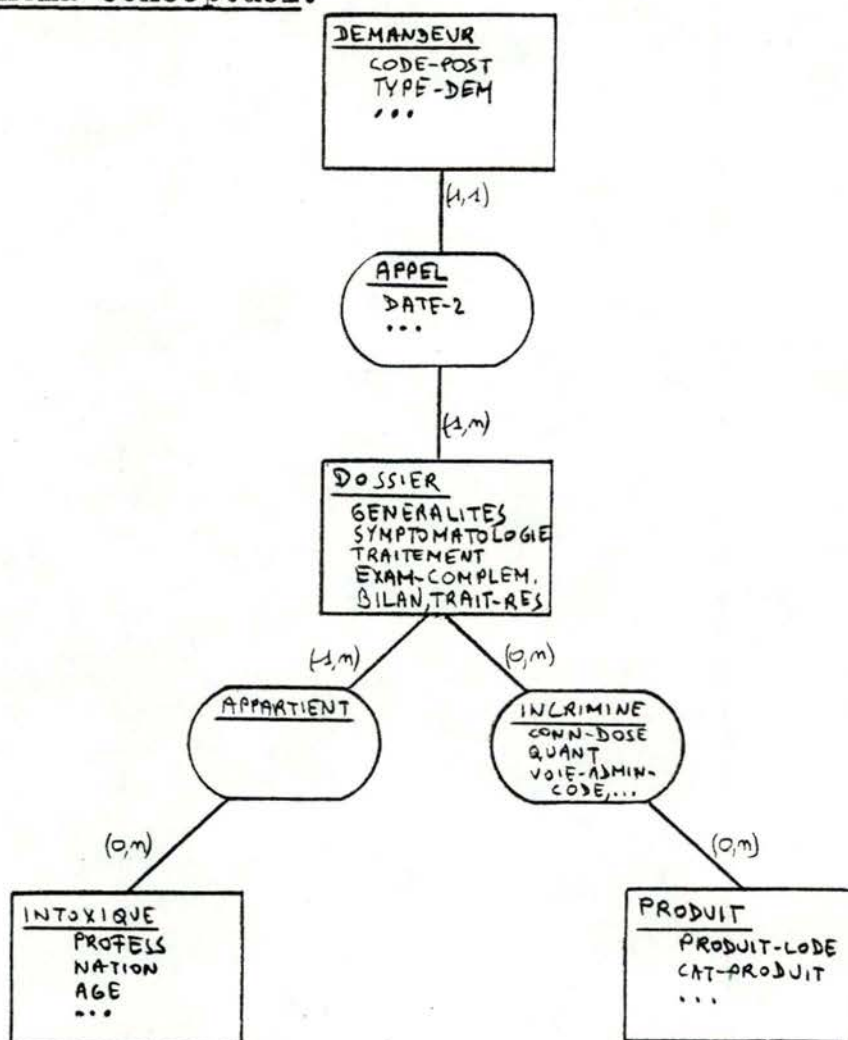
- NO-DOSSIER(NO-FICHE-MERE,DATE-1,LANGUE,CODE-MED,MOY-APPEL,URG-APPEL,NB-VICT,NB-VICT-DCD,CODE-INT)  
= art-généralités
- NO-DOSSIER(CODE-POST,TYPE-DEM,REGION,ORIG-DEM,MOTIF-DEM)  
= art-demandeur
- NO-DOSSIER(PROFESS,NATION,CAT-AGE,AGE,SEXE,POIDS,CAT-PODS,RACE,GROSS,ANTECEDS) = art-intoxiqué
- NO-DOSSIER(DATE-2,TYPE-EXPOS,TYPE-INTOX,LIEU-INTOX,APPEL-EXT)  
= art-circ-appel
- NO-DOSSIER(EX-SYMPT,RESPECT-CI,TRANSPORT,LIEU-TRAIT-PROP,LIEU-TRAIT-REEL,TRAITEM) = art-trait-res
- NO-DOSSIER(PRODUIT-CODE,CONN-DOSE,QUANT,VOIE-ADMIN-CODE,CAT-PRODUIT,CONN-COMPOS) = art-produit
- NO-DOSSIER(DEM-NOUV,REC-NOUV,ANALYSE,ENVOI-MED,RELATION,VALIDITE,EVOLUTION,INTERET,FICH-COMPL)  
= art-bilan
- NO-DOSSIER(SYMPTOMES) = art-symptom
- NO-DOSSIER(EEG,ECG,RECH-GROSS,RADIOGRAPHIE,ENDOSCOPIE,HISTOLOGIE,EXPLO-F-RESP,FOND-D-OEIL,EXAM-TOXICO,BIO-REIN,BIO-MIL-INT,BIO-SANG,BIO-HORM,AUTRES-EXAM)  
= art-exam-paracl
- NO-DOSSIER(ANTID-CHEL,ELIM-EPUR,NEUTR-ABSORP,TR-SYMPT,AUTRE-TRAIT)  
= art-traitement



### 2.4.3. Shéma conceptuel.

Pour mettre en lumière la méthode de recherche et la signification des différentes unités d'information que nous venons d'énumérer, nous présentons, dans ce paragraphe, le modèle de structuration logique de l'information (ou schéma conceptuel). A cette fin, nous utiliserons le modèle 'ENTITE-ASSOCIATION' exposé par F. Bobart (cfr. réf).

#### Shéma conceptuel.




---

F. BODART 'Problèmes d'organisation et méthodes d'analyse fonctionnelle.'

polycopie, Institut d'Informatique-FNDP-NAMUR.

#### 2.4.4. Description des articles et des fichiers logiques.

Il nous a paru intéressant de regrouper les différentes unités d'information pour en constituer trois articles logiques art-produit, art-dossier et art-complem. De cette manière, nous définissons trois types d'article correspondant, en réalité, à trois fréquences d'occurrence différentes.

00 F-DOSSIER.

01 FOR EACH NO-DOSSIER = art-dossier.

02 NO-DOSSIER.

02 DATE-1.

02 LANGUE.

02 CODE-MED.

02 MOY-APPEL.

02 URG-APPEL.

02 NB-VICT.

02 NB-VICT-DCD..

02 CODE-INT.

02 FICH-COMPL.

(suite)

02 CODE-POST.

02 APPEL-EXT.

02 TYPE-DEM.

02 EX-SYMP.

02 REGION.

02 RESPECT-CI..

02 ORIG-DEM.

02 TRANSPORT.

02 MOTIF-DEM.

02 LIEU-TRAIT-PROP.

02 PROFESS.

02 LIEU-TRAIT-REEL.

02 NATION.

02 TRAITEM.

02 CAT-AGE.

02 DEM-NOUV.

02 SEXE.

02 REC-NOUV.

02 AGE.

02 ANALYSE.

02 POIDS.

02 ENVOI-MED.

02 CAT-POIDS.

02 RELATION.

02 RACE.

02 VALIDITE.

02 GROSS.

02 EVOLUTION.

02 ANTECEDS.

02 INTERET.

02 DATE-2.

02 TYPE-EXPOS.

02 TYPE-INTOX.

02 LIEU-INTOX.



00 F-PRODUIT.

01 FOR EACH NO-DOSSIER = art-produit.  
02 NO-DOSSIER.  
02 PRODUIT-CODE.  
02 CONN-DOSE.  
02 QUANT.  
02 VOIE-ADMIN-CODE.  
02 TYPE-PROD.  
02 CONN-COMPOS.

00 F-COMPLEM.

01 FOR EACH NO-DOSSIER = art-complem.  
02 NO-DOSSIER.  
02 SYMPTOMES.  
03 HEPATO-DIG.  
04 SYMPT-CODE occurs 32 times.  
03 CUTAN-GENER occurs 40 times.  
04 SYMPT-CODE occurs 40 times.  
03 ENDOCRIN.  
04 SYMPT-CODE occurs 8 times.  
03 NEUROL-NEUROMUSC.  
04 SYMPT-CODE occurs 32 times.  
03 NEUROSENS.  
04 SYMPT-CODE occurs 40 times.  
03 RESPIR.  
04 SYMPT-CODE occurs 24 times.  
03 CARDIOVASC.  
04 SYMPT-CODE occurs 24 times.  
03 GENITO-URIN.  
04 SYMPT-CODE occurs 24 times.  
03 PSYCHIQ-AUTRES.  
04 SYMPT-CODE occurs 16 times.  
02 EEG.  
02 ECG.  
02 RECH-GROSS.  
02 RADIOGRAPHIE.  
02 ENDOSCOPIE.  
02 HISTOLOGIE.  
02 EXPLO-F-RESP.

02 FOND-D-OEIL.  
 02 EXAM-TOXICO.  
 02 BIO-REIN.  
 02 BIO-MIL-INT..  
 02 BIO-SANG..  
 02 BIO-HORM..  
 02 AUTRES-EXAM..  
 02 ANTID-CHEL..  
 02 ELIM-EPUR..  
 02 NEUTR-ABSORP..  
 02 TR-SYMPT..  
 02 AUTRE-TRAIT..

#### 2.4.5. Organisation physique des données.

Sur la figure de la page 151, nous présentons l'organisation physique des données en articles de longueur fixe. Trois types d'articles sont ainsi représentés, l'un à nombre d'occurrences par dossier fixe et unitaire, le second à nombre d'occurrences variable supérieur à zéro, le troisième enfin à nombre d'occurrences variable quelconque.

##### 2.4.5.a. Format interne.

Pour chaque donnée élémentaire, le format interne correspond au format défini dans le dictionnaire des données (traduit en code binaire et donc en bytes) sauf pour toutes celles notées de (\*) dans ce même dictionnaire. En ces cas, le code numérique décimal utilisé pour l'indexation des dictionnaires de (dé)codification est converti en un code interne binaire où chaque position (bit) de la zone physique de données correspond à une occurrence de ce type de donnée. Cette double codification nous est imposée par le fait que, pour ces types de données, plusieurs valeurs doivent pouvoir être prises en compte conjointement et que nous ne désirons manipuler ni des données de longueur variable, ni des groupes répétitifs à nombre d'occurrence variable.







#### 2.4.5.b. Format externe d'entrée.

Nous nous reportons aux deux fiches du dossier d'appel.

#### 2.4.5.c. Format externe de sortie.

Nous ne l'avons pas étudié en détail, préférant nous en tenir à quelques idées classiques et évidentes :

- édition (ou affichage) en clair, sous forme de tableaux de ventilation, de listes ou de textes bien structurés,
- possibilité d'une conversion en une représentation graphique plus parlante (notions de %, de cumuls, ...)

#### 2.4.6. Organisation physique des fichiers.

Nous aurions pu tenter une approche relative aux bases de données. Nous avons cependant préféré étudier et concevoir nous-même un modèle d'implémentation qui ne présente, au fond, pas de complexités. Aussi nous sommes-nous libéré des contraintes parfois lourdes de systèmes de gestion de bases de données. N'oublions pas non plus que la première de ces contraintes est de pouvoir en disposer, la seconde de maîtriser ses possibilités pour pouvoir en extraire un maximum de facilités au plan théorique. D'un autre côté, tout S.G.B.D. possède ses caractéristiques, si différentes et parfois si spécifiques qu'il nous faudrait trop de temps pour les comprendre, les évaluer, les comparer et les exploiter au mieux.

L'impératif qui nous est posé consiste à privilégier un certain nombre d'accès. Tandis que les programmes SINBAD (cfr. annexe 4) tentaient de réaliser l'accès isotrope, tout en le justifiant bien évidemment, les concepteurs du système lyonnais donnaient leurs préférences à cinq types d'accès :

- accès par le produit,
- accès par la symptomatologie,
- accès par la classe d'âge,
- accès par l'évolution de l'intoxication,
- accès par le type d'intoxication.



A cet éventail, les responsables du centre de Bruxelles désirent ajouter :

- accès par la date,
- accès par le traitement,
- accès par la région.

Le principe sous-jacent à la création de fichiers inversés dans le but de privilégier certains critères d'interrogation, consiste à rassembler, dès l'enregistrement des données, les identifiants d'articles qui répondent à ces critères. En un minimum de transferts, les programmes d'application extrayent ces identifiants avant d'aller rechercher, grâce à eux, les articles ou parties d'articles désirés.

#### 2.4.6.a. 1ère possibilité : organisation type-1.

Toujours dans cette optique d'économie de transferts, nous éclaterons chaque enregistrement en différents articles, correspondants aux différentes unités d'information recensées. Un tel éclatement nous permettra d'obtenir des articles de faible longueur et donc un nombre élevé d'articles du même type par page.

Sans privilégier outre mesure certains critères, les programmes consulteront la totalité des parties de dossiers visées par la requête de l'utilisateur. Pour chacune des parties retenues, le NO-DOSSIER établira le lien, si nécessaire, avec les autres éléments du même dossier, de manière à chercher les éléments demandés.

Ce procédé s'applique sans peine aux parties d'enregistrement dont le nombre d'occurrences et la longueur sont fixes. Nous ne pourrions donc traiter de cette façon :

- art-produit,
- art-traitement,
- art-symptom,
- art-exam-paracl.

L'accès à la fiche complémentaire est subordonné à l'accès à la rubrique art-bilan. Nous y placerons donc un pointeur, déterminant, soit l'absence de cette fiche (pointeur = 0),

= 17 BYTES

TRAIT-RES

3	1	4	1	2	6	3
NO-DOSSIER	EX-SYMP	RESPECT-CI	TRANSPORT	LIEU-TRAIT	TRAITEM	NO-étude

= 20 BYTES

PRODUIT

3	3	1	6	1	1	1	4
NO-DOSSIER	PRODUIT-CODE	CONN-DOSE	QUANT	TYPE-ROLE-ADMIN	CONN-COMPOS	pointeur	page dépl.

= 17 BYTES

BILAN

3	1	1	6	1	1	1	1	1	1	1	1
NO-DOSSIER	DEM-NOUV	REL-NOUV	ANALYSE	ANAL-NO	LIEU	ENVOI-MED	RELATION	VALIDITE	EVOLUTION	INTERET	FICH-COMPL

= 37 BYTES

SYMPTOMATOLOGIE

3	4	5	1	4	5	3	3	3	2	4
NO-DOSSIER	troubles hépatodigestifs	signes cutanés et généraux	signes endocr.	signes neurolog. et neuromus.	signes neurosens.	signes respir.	signes cardio-vasc.	signes gastro-intest. urin.	sign. psych. et autres	pointeur

= 34 BYTES

EXAM-PARAACL

3	1	1	1	3	4	2	1	1	2	2	2	5	4
NO-DOSSIER	EEG	EEG	RECH-GROSS	RADIOGRAPHIE	ENDOSCOPIE	HISTOLOGIE	EXAM-TOXICOL	BIO-BIO-REIN	BIO-MILINT	BIO-SANG	BIO-HORM	AUTRES-EXAM	pointeur





soit son emplacement exact.

Néanmoins, toute recherche peut directement s'appliquer sur l'ensemble des articles art-symptom, art-traitement ou art-exam-paracl, quitte à devoir revenir sur un autre article du dossier via le NO-DOSSIER.

Afin qu'un seul pointeur suffise à l'accession à ces trois types d'articles, articles de longueur différente, nous ne placerons pas d'adresse mais un élément du calcul d'adresse, à savoir le numéro d'ordre de cet article dans ces fichiers. L'adresse sera alors obtenue en multipliant ce numéro d'ordre par la longueur du type d'article du fichier considéré.

Les modifications opérables lors d'une phase de mise à jour doivent pouvoir consister en l'adjonction d'une fiche complémentaire. Nous établirons donc, à l'intérieur de chacun des fichiers F-TRAITEMENT, F-SYMPATOM et F-EXAM-PARACL, un chaînage d'articles à l'aide de pointeurs d'adresse situés en bout d'article.

TABLE DES LONGUEURS

SYMPTOMAT.	$\alpha$

F-TRAIT-RES

NO-DOSS		no ordre

MODULE  
DE  
CALCUL  
D'ADRESSE

F-SYMPATOMOLOGIE

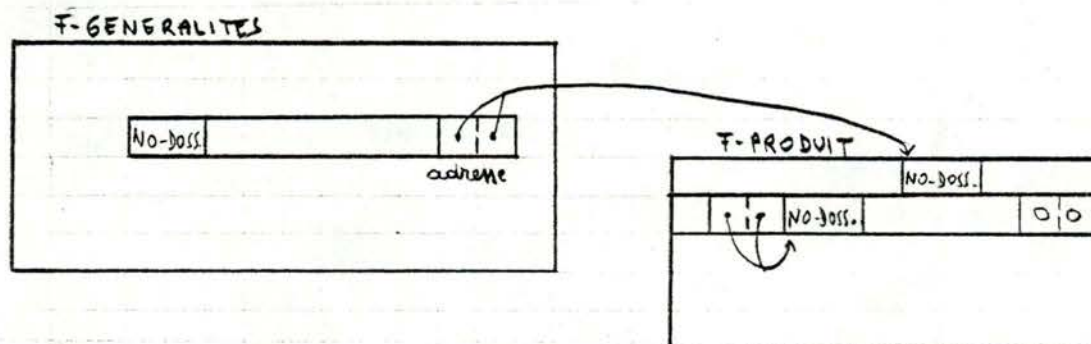
NO-DOSS			
NO-DOSS			
NO-DOSS			0

Pareille organisation laisse donc la possibilité d'incorporer, à tout moment, une n-ième fiche complémentaire (donc un article de chacun des trois types), en sauvegardant le lien d'appartenance au dossier et en intégrant la notion de bilan évolutif (chaînage).



Quant à l'implémentation des articles art-produit, caractérisés eux-aussi par un nombre d'occurrences variable, nous procéderons d'une façon analogue. Un pointeur d'adresse (n° page, dépl.) incorporé à l'article art-généralités indique la position de l'article art-produit correspondant (éventuellement inconnu, pointeur = 0).

Si plusieurs articles art-produit répondent du même dossier, ils seront également chaînés.

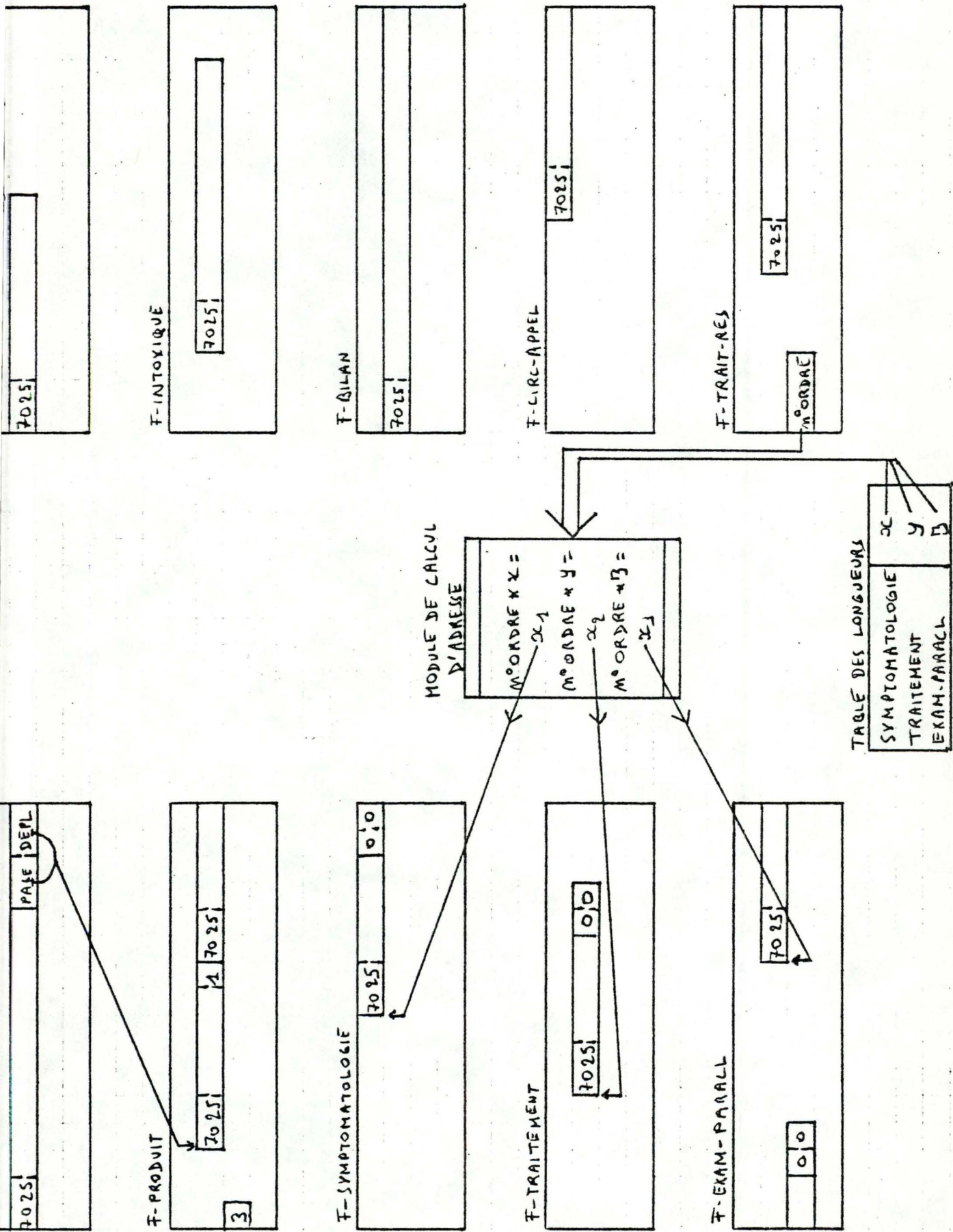


Toute modification relative à une donnée enregistrée n'altère pas la structure physique des fichiers. Lorsqu'elle porte sur des données non-enregistrées, elle est supportée par l'organisation en listes d'articles.

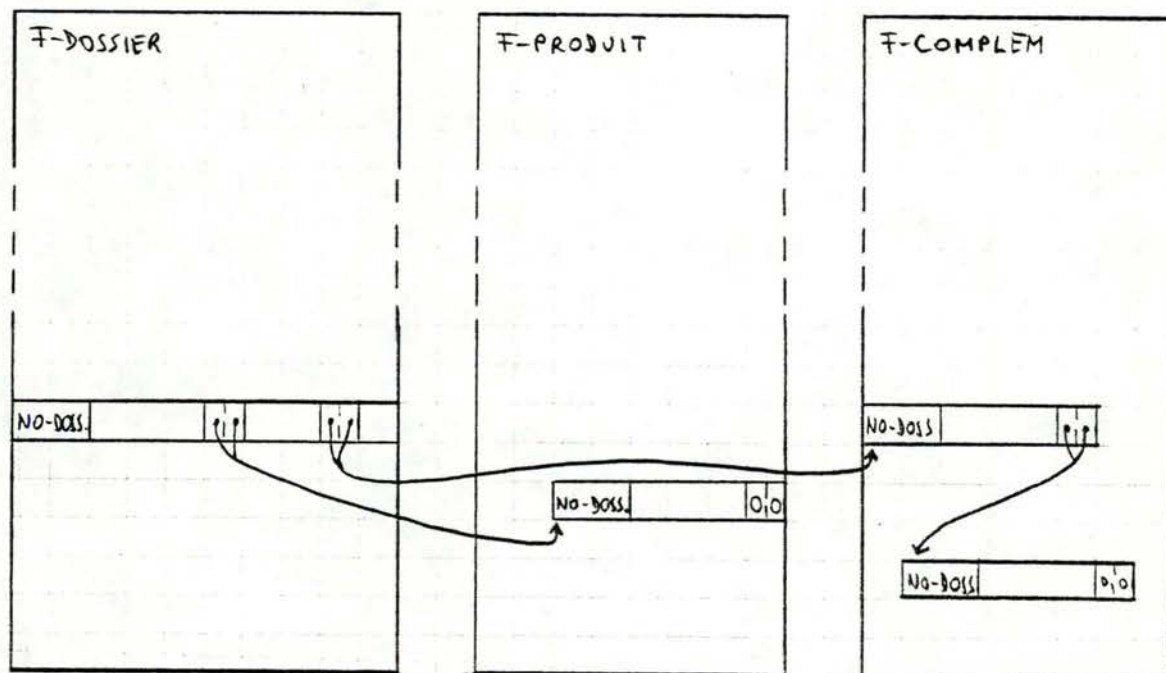
Considérant toutes ces remarques, nous pouvons schématiser la structure globale des différents fichiers (page suivante). Basée sur une structure matricielle et sur la redondance de l'indicatif, elle présente l'avantage d'être assez simple et, nous l'espérons, suffisamment performante en ce qui concerne la place mémoire occupée et le temps de réponse moyen.

#### 2.4.6.b. Deuxième possibilité : l'organis. séquentielle.

L'organisation séquentielle que nous testerons également consiste à créer un type de fichier par nombre moyen d'occurrence des types d'article. De cette manière, nous recensons trois types de fichiers ainsi schématisés :







A l'intérieur de chacun de ces fichiers, la structuration est rigoureusement semblable à celle de l'organisation type-1.

Remarques : - On travaille toujours en longueur fixe.  
 - Il existe toujours un seul référentiel.  
 - Toute recherche peut porter sur n'importe quel fichier, le NO-DOSSIER permettant de retrouver les articles d'autres types associés.

#### 2.4.6.c. Troisième possibilité : organisation séquentielle avec fichiers inversés.

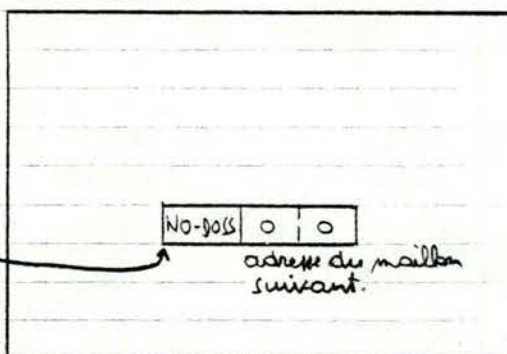
Ce modèle d'organisation présente les mêmes caractéristiques que celui présenté au point b, si l'on excepte les fichiers inversés. En effet, nous privilégierons les huit critères d'interrogation retenus par les responsables du centre, à savoir, le produit, le traitement, la symptomatologie, l'évolution, la région, l'âge, la date et le type d'intoxication, ce qui entraînera la création de huit fichiers inversés.

Chacun de ces fichiers sera implémenté classiquement comme suit :

TABLE-INV-REGION

REG.	adresse	REG.	adresse
15			

F-INV-REGION



#### 2.4.7. Evaluation des performances globales.

La méthode que nous utilisons consiste à partir des trois modèles de structure d'organisation que nous avons définis, pour en calculer les performances en fonction de différentes hypothèses et d'un jeu de paramètres plus ou moins étoffé.

##### 2.4.7.a. Les modélisations.

Nous terminons de les présenter au point 2.4.6.

##### 2.4.7.b. Les quantifications.

Certains paramètres sont communs aux trois modèles. Ce sont :

- nombre actuel de dossiers :  $100\ 000 < \text{NBDOS} < 120\ 000$
- évolution annuelle :  $+ 5\ \% / \text{an}$
- nombre de dossiers en 78 :  $21\ 000$
- nombre moyen (par dossier) de cartes produits :  
 $1 < \text{NBCPRO} < 1,2$
- nombre moyen de fiches complém. :  $0,2 < \text{NBFCPL} < 0,4$

D'autres paramètres sont fonctions de chaque modèle de structure d'organisation.



- tableau des paramètres relatifs à l'organisation type-1.

GENERAL	DEMAND	INTOXIQ	CIRC- APPEL	TRAIT- RES	BILAN	PRODUIT	EXAM- PAR	SYMPTOM	TRAITEM	TOTAL
NBDOSS = 110 000 LGPAGE = 2 000										
(1) longueur des articles en nombre de bytes.	24	12	18	14	17	17	20	34	31	220
(2) nombre d'articles moyen par dossier.	1	1	1	1	1	1	1,1	0,3	0,3	
(3) longueur moy. des art. = (1) * (2)	24	12	18	14	17	17	22	10,2	9,3	150,6
(4) volume total occupé par fichier										
= (3) * NBDOSS	2640	1320	1980	1540	1870	1870	2420	1122	1023	16546
(5) nombre d'articles par page = LGPAGE / (1)	83	166	111	142	117	117	100	58	64	
(6) nombre de pages occup. = (NBDOSS/(5)) * (2)	1326	663	991	775	941	941	1210	569	516	8323

## - tableau des paramètres relatifs à l'organisation type-2.

NBDOSS = 110 000 LGPAGE = 2 000	DOSSIE	PRODUIT	COMPLE	TOTAL
(1) longueur des articles	87	20	88	191
(2) nombre moyen d'art. par dossier	1	1,1	0,3	
(3) longueur moy. des art.	87	22	26,4	131,4
(4) volume total occupé par fichier	9570	2420	2904	14454
(5) nombre d'articles par page	22	100	22	
(6) nombre de pages occup.	5000	1210	1500	7294

## - paramètres relatifs à l'organisation type-3.

La quantification de l'organisation type-2 reste valable. Il nous faut cependant y ajouter tout ce qui relève de l'implémentation des huit fichiers inversés.

b.1. Région.

1 table de  $100 \times (1+4)$  = 500 bytes max.  
 110 000 pointeurs de  $(3+4)$  = 770 000 bytes max.  


---

 770 500 bytes max.  
 500 000 bytes moy.

b.2. Type d'intoxication.

1 table de  $12 \times (1+4)$  bytes = 60 bytes  
 110 000 pointeurs de  $(3+4)$  = 770 000 bytes max.  


---

 770 060 bytes max.  
 500 000 bytes moy.

Remarque: La taille de la table nous semblant négligeable, nous n'en tiendrons désormais plus compte.

b.3. Date.

Point n'est besoin de créer un fichier inversé complet, car nous pouvons nous baser sur les NO-DOSSIER qui sont attribués séquentiellement. Il nous suffit de disposer d'une petite table de points de repère qui nous permette de localiser, pour chaque



mois, la bonne fourchette de NO-DOSSIER.

Cette table peut aussi être considérée de taille négligeable.

Pour mémoire :  $(12 * NBANNEES * qq \text{ bytes})$

#### b.4. Evolution.

1 table = (pour mémoire)

110 000 pointeurs de  $(3+4)$  = 770 000 bytes

De l'avis des responsables du centre, seule la distinction 'évolution au décès' ou 'évolution au non-décès' importe. De ce fait, une seule liste de moins de 100 000 bytes suffit amplement.

#### b.5. Age.

1 table = (pour mémoire)

110 000 pointeurs de  $(3+4)$  = 770 000 bytes max.

650 000 bytes moy.

#### b.6. Produit.

La table pourrait être avantageusement combinée avec une table de (dé)codification du catalogue des produits, l'article devenant ainsi constitué :

PRODUIT-CODE	NOM-PRODUIT	adresse du premier maillon du fichier inversé
--------------	-------------	---

1 table de 20 000 \* 4 bytes = 80 000 bytes

110 000 pointeurs de  $(3+4)$  = 770 000 bytes

850 000 bytes

\*1,1= 935 000 bytes

#### b.7. Symptomatologie.

Nous appliquons le même principe que ci-dessus.

1 table de 200 \* 4 bytes = 800 bytes (négl.)

110 000 pointeurs  $(3+4)$  = 770 000 bytes

\*0,3= 231 000 bytes

Un même dossier pouvant se trouver dans plusieurs listes, nous estimerons la place nécessaire à 1 000 000 bytes.

#### b.8. Traitement.

1 table = (pour mémoire)

110 000 pointeurs (3+4+1) = 880 000 bytes

\* 0,3 = 264 000 bytes

même remarque que b.7. = 792 000 bytes

Lorsque nous opérons la somme de ces différentes allocations de mémoire, nous approchons les 4,5 MB, ce qui nous paraît fort important vis-à-vis des 14,5 MB qu'occupe l'ensemble des fichiers directs.

#### 2.4.7.c. Caractéristiques physiques des mémoires de masse.

Cet aspect peut à lui seul constituer le facteur décisif en faveur de l'un ou l'autre type d'organisation car c'est essentiellement à partir de ces caractéristiques physiques, parfois fort différentes d'un matériel à l'autre, que se base l'essentiel de nos évaluations et, par là-même, de notre appréciation globale.

Conscients de ce fait, nous aurions pu tester les performances de nos trois modèles sur différents matériels. Nous ne l'avons pas réalisé, faute de temps et surtout désireux de nous éviter, au bout du compte, un choix impossible face à des résultats peut-être contradictoires.

Nous prendrons donc comme référence un modèle de disque inexistant, mais dont les spécifications physiques nous semblent correspondre au niveau technique actuel.

- capacité totale	80 MB
- nombre de cylindres	400
- nombre de pistes / cylindre	20
- nombre de pages (blocs) / piste	5
- taille fixe d'une page	2000 bytes
- taux de remplissage d'une page	100 %
- temps de transfert d'une page	5 ms
- temps d'accès à une page du cylindre	8 ms
- temps d'accès moyen	30 ms
- temps d'accès maximum	50 ms



2.4.7.d. Projet.

Nous nous efforçons d'évaluer, sur base d'hypothèses que nous estimons raisonnables, le temps de réponse que délivrerait chacun des trois modèles d'implémentation présentés, au niveau de la consultation des fichiers, étant entendu qu'elle constitue le point d'intérêt crucial de l'ensemble du sous-système. Pour juger de l'efficacité des différentes possibilités, nous nous sommes proposés de considérer un échantillon représentatif des types de question susceptibles d'être couramment formulées par les utilisateurs.

type	critères de sélection	fréquence
Q1	produit	0,3
Q2	type-intoxication, région, âge	0,05
Q3	produit, traitement, évolution	0,2
Q4	produit, traitement, évolution, symptomatologie, région	0,2
Q5	type-intoxication, région, un critère non-privilegié(a)	0,15
Q6	type-intoxication, région, deux critères non-privilegiés(a et b), âge, évolution	0,05
Q7	un seul critère non-privilegié(b)	0,05

Afin d'éviter la multiplicité des combinaisons, due au nombre d'hypothèses posées, nous nous en tiendrons délibérément à ces sept types de requêtes et à la pondération y attendant.

L'ordre de prise en compte des critères dépend en réalité de chaque type d'organisation physique des fichiers. Les critères choisis pour évaluer les performances des systèmes seront donc numérotés d'une manière unique pour chaque type de requête. Ainsi,

- Q1 - 1.produit
- Q2 - 1.âge / 2.région / 3.type-intox
- Q3 - 1.produit / 2.évolution / 3.traitement
- Q4 - 1.produit / 2.région / 3.évolution / 4.symptom /

## 5. traitement

- Q5 - 1.région / 2.type-intox / 3.a
- Q6 - 1.âge / 2.région / 3.type-intox / 4.évolution /  
5.b / 6.a
- Q7 - 1.b

La détermination de cet ordre de prise en compte sera fixée, pour chaque requête de l'utilisateur, par un module d'optimisation, dont les principaux critères de jugement seront :

- l'existence d'un fichier inversé pour le critère considéré,
- la longueur du fichier à dérouler,
- la proportion d'éléments (dossiers) qui devraient satisfaire ce critère, calculée sur une loi de distribution approchée.

2.4.7.e. Hypothèses.

- \* Nous n'admettons qu'un utilisateur (un seul terminal).
- \* L'ensemble des fichiers n'occupe qu'un seul disque.
- \* Nous nous en tenons, en ce qui concerne la phase de calcul, au temps de travail de l'unité-disque, puisqu'elle constitue le goulot d'étranglement du système.
- \* Nous supposerons également que, sur chacun des critères évoqués au point d,
  - soit la distribution du nombre de dossier est uniforme (Ha),
  - soit la valeur précisée dans la requête est trois fois plus qu'équiprobable (Hb),
  - soit la valeur est une fourchette qui sélectionne le dixième des dossiers encore admissibles (Hc).
- \* Nous admettrons enfin que chaque cylindre est exclusivement occupé par un type de fichier, de façon à éviter les changements de cylindre intempestifs.

2.4.7.f. Estimation des temps.

En prélude aux calculs, nous présentons les fréquences des critères, relatives aux trois hypothèses de distribution que nous avons envisagées.



critère	Ha	Hb	Hc
âge (31+35+97)=163 (cfr. remarque)	0.006	0.018	0.10
type-intox = 12	0.08	0.25	0.10
produit = 20 000	0.00005	0.00015	0.10
symptom = 200	0.005	0.015	0.10
traitem = 46	0.02	0.06	0.10
évolution = 4	0.25	0.75	0.10
région = 100	0.01	0.03	0.10
a = 50	0.02	0.06	0.10
b = 100	0.01	0.03	0.10

rem.:  
 1 mois =  
 31 jours  
 3 ans =  
 35 mois+  
 31 jours  
 100 ans=  
 97 ans+  
 3 mois

Exemple : - Soit Q1 pour l'organisation type-1.

Le fichier F-PRODUIT occupe 1210 pages concentrées sur 13 cylindres..

Le temps de réponse est donc équivalent à la somme de :

1197 accès de 8 ms	= 9576 ms
13 accès aléatoires	= 390 ms
1210 transferts de page	= 6050 ms
	<hr/> 16016 ms

Selon Ha, la recherche débouche sur une liste de 6 dossiers, selon Hb, de 16.5 dossiers et selon Hc, de 11000 dossiers..

La prise en compte d'autres critères devrait porter sur ces chiffres.

- Soit Q1 pour l'organisation type-2.

Les résultats sont identiques à ceux fournis par l'organisation type-1 puisque le fichier F-PRODUIT a une structure identique dans ces deux modèles d'organisation.

- Soit Q1 pour l'organisation type-3.

Le fichier inversé des produits s'étale sur 468 pages. Il nous permet d'obtenir la liste demandée en un maximum de :

5 accès aléatoires	= 150 ms
463 accès dans le cylindre	= 3704 ms
468 transferts de page	= 2340 ms
	<hr/> 6194 ms

Nous ne détaillerons pas ici l'ensemble des calculs opérés. Les résultats sont néanmoins présentés sous la forme d'un tableau synthétique en annexe 6.

#### 2.4.7.g. Premier commentaire.

Les résultats relatifs aux trois types d'organisation présentés appellent quelques remarques.

Alors que les chiffres calculés ne représentent que le temps d'accès disque nécessaire à la prise en compte des seuls éléments donnés (critères de recherche), le résultat obtenu par l'organisation type-2 se révèle totalement inacceptable (42 secondes !). Son seul intérêt réside dans la possibilité de le comparer aux deux autres modèles.

L'organisation type-1, que nous avons imaginée, assez intéressante sur le plan de l'ensemble car elle permet de niveler les temps de réponse, pêche, elle-aussi, par son résultat global ( $18 \text{ sec.} < X < 33 \text{ sec.}$ ), principalement dans la perspective d'évolution du nombre de dossiers à exploiter.

L'organisation type-3 présente, quant à elle, une moyenne alléchante (12 secondes) mais pénalise trop tout critère non privilégié (66 secondes). Il est vrai que ce temps reste sensiblement équivalent quel que soit le nombre de critères non-privilégiés à considérer, à l'instar de l'organisation type-1, pour laquelle il nous faut cumuler les temps de recherche si les critères de sélection de dossiers sont disséminés dans des fichiers différents.

Cette dernière remarque est à considérer avec attention, surtout si l'on sort du cadre de la consultation au profit de celui des recherches statistiques générales.

#### 2.4.7.h. Prise en compte du temps de réponse global (accès disque).

Si nous considérons, en outre, le temps nécessaire au trai-



tement des éléments demandés, toujours au niveau de la consultation 'on-line', il apparaît que trois facteurs entrent en ligne de compte. Ce sont :

- le nombre d'éléments demandés,
- leur appartenance à un seul fichier ou à des fichiers différents,
- le nombre de dossiers ayant satisfait au filtrage de la question (éléments donnés).

#### h.1. Un seul élément demandé.

Quelle que soit l'organisation choisie, le temps d'accès est pratiquement identique si le nombre de dossiers sélectionnés est faible ( $\leq 10$ ).

Si ,par contre, le nombre de dossiers sélectionnés dépasse ce chiffre, l'avantage revient à l'organisation type-1, car la recherche nécessite, pour un fichier moins étendu, moins de changements de cylindre.

Cet avantage est loin d'être décisif puisque, pour cent dossiers comme pour mille, il avoisine la seconde.

NBDOSS	ORG. TYPE-1	ORG. TYPE-2 ou 3
100	1.520 sec	2.400 sec
1000	13.308 sec	14.100 sec

#### h.2. Plusieurs éléments demandés.

C'est à ce niveau que joue l'appartenance à un seul ou à différents fichiers. De manière générale, tout accès à un art-produit, art-symptom, art-trait ou art-exam-paracl nécessite un accès préalable, soit au fichier F-DOSSIER (org. type-2 ou 3), soit aux fichiers F-TRAIT-RES ou F-GENERALITES (org. type-1).

Quant aux éléments intégrés dans les autres articles, nous pouvons en déduire :

- que l'organisation type-1 privilégie les questions dans lesquelles les éléments demandés sont peu nombreux et concentrés dans peu de fichiers différents, quel que soit finalement le nombre de

dossiers encore valides.

- que l'organisation type-2 ou 3 privilégie les questions dans lesquelles le nombre d'éléments demandés importe peu, mais pour lesquelles le nombre de dossiers encore valide est faible.

Il nous paraît difficile de fixer des limites précises à ces termes 'faible' ou 'peu nombreux', tant il est totalement aléatoire de prévoir le nombre et la combinaison des éléments susceptibles d'être demandés, ainsi que la fréquence de chaque type de question.

D'autre part, nous ne pouvons perdre de vue l'ensemble des recherches statistiques régulièrement entreprises sur l'éventail des dossiers, recherches qui, présentant des exigences de temps nettement moins impératives, constituent à elles seules la moitié de l'ensemble des requêtes adressées au système.

A titre indicatif, nous présentons ci-dessous quelques chiffres, à interpréter.

NBD OSS	NBCRIT	répartition des critères	TEMPS D'ACCES (sec)		
			ORG.-1	ORG.-2	-3
100	1	1 / 0 / 0	1.498	2.400	
100	3	1 / 1 / 1	4.494	2.400	
100	3	2 / 1 / 0	2.996	2.400	
100	3	3 / 0 / 0	1.498	2.400	
			<u>moy 2.996</u>	<u>moy 2.400</u>	
1000	1	1 / 0 / 0	11.898	14.100	
1000	3	1 / 1 / 1	35.694	14.100	
1000	3	2 / 1 / 0	23.796	14.100	
1000	3	3 / 0 / 0	11.898	14.100	
			<u>moy 23.796</u>	<u>moy 14.100</u>	

Rem. : La moyenne ne porte que sur les cas où trois critères entrent en ligne de compte.

Une fois de plus, ces chiffres révèlent, dans les rares cas où l'organisation type-1 se trouve à son avantage, une dif-



férence non-significative.

En moyenne, l'organisation matricielle simple présente des résultats plus concluants et surtout une valeur maximum tout à fait acceptable.

#### 2.4.7.i. Choix d'une organisation définitive.

##### Conclusion.

Au terme de cette étude de performance réalisée malgré la précarité de certaines estimations, il nous faut déterminer le type d'organisation à implémenter.

Au vu des résultats avancés et des commentaires qui en ont découlé, le lecteur aura, sans aucun doute, deviné le choix qui s'imposait.

L'organisation type-3 (fichiers inversés) recueille donc notre suffrage, principalement grâce au temps de réponse qu'elle garantit. Malgré les quelques MB supplémentaires dont nous devons payer le prix, ce choix se révèle indispensable à la réalisation des objectifs qui nous étaient impartis.

#### 2.5. Analyse des traitements.

##### 2.5.1. Enregistrement.

En permanence, différents pointeurs retiennent, pour chaque fichier, l'adresse de la première position disponible. Chaque enregistrement est éclaté en différents articles qui prennent place séquentiellement dans chacun des fichiers concernés. La mise à jour des différents pointeurs de place libre s'opère en même temps que se garnissent les deux zones d'adresse et les pointeurs de art-dossier.

##### 2.5.2. Mise\_à\_jour.

Toute donnée entrée doit pouvoir être corrigée. La fonction de mise à jour, sélectionnable en un mode-utilisateur privilégié (sous le contrôle de mots de passe), retrouve l'article erroné (cfr. 2.5.4.) avant de le mettre à jour ou le corriger.

Les modifications seront en fait peu nombreuses si les res-

ponsables du système décident de n'enregistrer les dossiers que plusieurs jours après l'appel.

Si l'enregistrement 'on-line' ne se justifie pas, la mise à jour, de son côté, se fera de manière plus aisée et plus sûre si elle peut s'exécuter en mode conversationnel. Cependant, rappelons une fois encore notre position; cette fonction ne peut se trouver que sous le contrôle exclusif d'une équipe restreinte. Il ne faut pas que chaque utilisateur puisse modifier à sa guise certaines données, sous peine de voir apparaître des erreurs là où il n'en existait pas, en sus de celles qui existaient réellement.

### 2.5.3. Suppression.

Nous ne trouvons pas de réel motif à l'emploi de cette fonction. Tout dossier enregistré doit le rester (fichier historique). En cas d'erreur, il peut à la rigueur être modifié mais la place attribuée aux articles qui le composent reste fixe, du moins pour les articles à nombre d'occurrences fixe.

On pourrait penser qu'il est inutile de conserver des dossiers d'intérêt mineur. Il ne faut cependant pas oublier que, sous l'angle purement statistique, aucun cas ne peut être écarté.

D'un autre côté, personne ne peut juger a priori du degré d'intérêt d'un cas. Il s'est en effet quelquefois avéré qu'une accumulation de cas 'banaux' entraînait une prise de conscience d'un problème que jamais personne n'aurait pu détecter autrement.

### 2.5.4. Interrogation. Consultation.

Bien évidemment, il s'agit là de la fonction la plus intéressante du point de vue du système global. Tout doit être fait pour faciliter le travail de l'utilisateur, car c'est à ce niveau que les performances de l'outil seront surtout évaluées.

Nous avons recensé ci-après la liste des types de requête susceptibles d'être posées :



- consultation d'un dossier sur base de son numéro, en tout ou en partie,

- sélection de numéros de dossiers sur base d'un critère ou d'une combinaison de plusieurs critères à l'aide d'opérateurs logiques (ET, OU, SAUF, NON) et arithmétiques ( $<$ ,  $>$ ,  $=$ ,  $\leq$ ,  $\geq$ ,  $\neq$ ), ces critères étant principalement choisis parmi ceux que nous avons privilégiés,

- établissement de séries de fréquences, sur base d'une liste de numéros de dossiers (obtenus par exemple au point précédent), fréquences d'apparition des produits ou de tout autre donnée.

Ces trois points, éventuellement combinés, recouvrent à peu près la totalité des requêtes envisageables sur un ensemble de dossiers. Il va sans dire que l'enchaînement de ces différents points doit pouvoir s'établir de manière transparente pour l'utilisateur, mais, du point de vue de la définition des traitements, nous pouvons les dissocier sans aucun problème.

Examinons tour à tour chacun d'entre eux.

#### 2.5.4.a. Consultation d'un dossier sur base de son n°.

Tout numéro de dossier est constitué des deux derniers chiffres du millésime et de son numéro d'ordre dans l'année, codé en cinq chiffres. Il serait à tout le moins intéressant d'user de ce NO-DOSSIER en tant qu'élément du calcul d'adresse. Seule la réinitialisation en début de chaque année empêche l'adressage direct à l'aide de ce numéro ( $\text{NO-DOSSIER} * \text{longueur de l'article} = \text{adresse virtuelle}$ ). Pour pallier à cet inconvénient, nous constituons une table intermédiaire (table-adr) à laquelle nous accéderons à l'aide des deux premiers chiffres du NO-DOSSIER. L'entrée correspondante, cumulée au numéro d'ordre du dossier, délivrera le numéro d'ordre absolu, indépendant des années. Pondérée par la longueur des articles (table des longueurs), ce numéro d'ordre détermine l'adresse exacte de l'article du dossier recherché. Ce mécanisme devient inutilisable pour les fichiers F-PRODUIT et F-COMPLEM. C'est donc à dessein que nous avons placé dans art-dossier l'adresse éventuelle de l'article art-produit cor-

respondant dans F-PRODUIT ainsi qu'un pointeur, qui est en réalité un numéro d'ordre à multiplier par la longueur de l'article de F-COMPLEM.

2.5.4.b. Requêtes spécifiques sur base de critères définis par l'utilisateur.

C'est à ce niveau qu'il nous faut justifier le mode d'implémentation que nous avons retenu. Si la technique des fichiers inversés permet d'obtenir directement, sans recherche comparative, la liste des indicatifs répondant au critère désiré, au prix d'une implantation en mémoire de nombreuses et parfois longues listes de NO-DOSSIER, et de minimiser ainsi le nombre d'accès en mémoire auxiliaire, cela se passe aux dépens de tout autre critère non-privilegié.

Notre étude de performance l'a démontré; dans l'organisation que nous avons sélectionnée, l'interrogation d'un fichier selon un critère quelconque non-privilegié de l'article art-dossier exige un minimum d'une minute de temps de réponse.

Pour parer à cet inconvénient, nous recommanderons à l'utilisateur d'introduire au préalable dans sa question l'un des critères privilégiés, pour, à la fois, limiter le champ des investigations et le temps de réponse.

C'est d'ailleurs pour cette raison qu'il nous a paru utile de définir huit critères privilégiés, soit trois de plus qu'à Lyon. L'expérience montrera peut-être l'inutilité de certains d'entre eux, ou la nécessité de privilégier d'autres critères encore.

Nous présentons ci-dessous l'algorithme de sélection simple (sans opérateurs logiques).

Dans le cas de critères non-privilegiés, il est intéressant de traiter en une seule étape, lors du même transfert, ceux qui se situent dans un même type d'article et donc dans un même fichier.

L'algorithme admet alors comme paramètres, outre la LISTE-X, un ensemble de critères de sélection. Le fichier est ensuite déroulé de façon à ce que chaque article de NO-DOSSIER

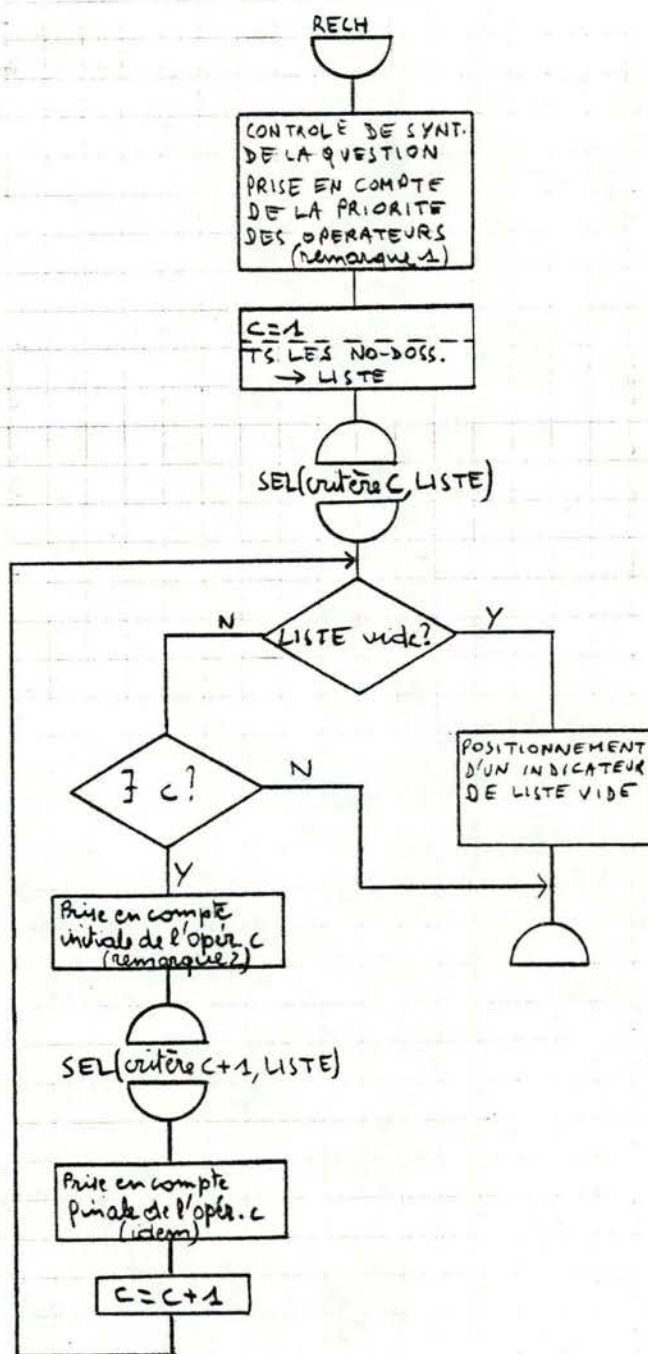


acceptable (LISTE-X d'entrée) soit testé par rapport aux valeurs des critères de sélection définis.



Si nous considérons le niveau des opérateurs et donc des combinaisons entre critères, l'algorithme de recherche applicable s'enrichit de la manière suivante :

Cet algorithme de recherche traite les éléments donnés de la requêtes, de manière à constituer une liste de NO-DOSSIER qui servira de base au traitement des éléments recherchés. Il peut



donc, selon les désirs formulés par l'utilisateur,

- soit afficher ces numéros de dossiers et déclencher la fonction de consultation,
- soit déclencher la fonction développée au point 2.5.4.c.

Remarques : -1- La notion d'optimisation répond à un fait bien précis. Lors des recherches opérées sur les fichiers, il est préférable de traiter en priorité le critère le plus sélectif, de façon à limiter le plus tôt possible le champ des investigations. La sélectivité des critères peut s'étudier par quelques études statistiques générales préalables.

-2- Chaque opérateur inclut un traitement différent. Si l'on considère les opérateurs ET et SAUF, il est évident que la recherche se prolonge sur base de la dernière liste de NO-DOSSIER produite (LISTE-X). L'opérateur OU implique au contraire une recherche non-limitée à cette liste. Il faut donc oublier la liste actuelle pendant l'appel de la routine SEL, puis opérer la fusion ou l'exclusion des éléments figurant ou ne figurant pas dans ces listes (prise en compte finale).

#### 2.5.4.c. Etablissement de séries de fréquences.

Cette fonction traite la liste des NO-DOSSIER obtenus après la prise en compte d'un ou de plusieurs critères de recherche, consulte les fichiers adéquats et transfère les éléments nécessaires en mémoire centrale. Ces éléments peuvent alors subir des traitements statistiques (calcul de moyenne, ajustement d'une loi de distribution, ventilation simple ou double selon différentes données, ...) avant d'être envoyés à l'utilisateur. Il est inutile de détailler ici toute théorie statistique à ce sujet, là n'est pas notre tâche.



### 2.5.5. Production\_d'états\_statistiques.

Combinant à la fois les points b et c du paragraphe précédent, cette fonction ne se distingue de l'interrogation que par le mode sur lequel elle se sélectionne. Elle est essentiellement destinée à des études plus complètes, d'envergure a priori plus large que la précédente, et ne nécessite pas de réponse rapide. Nous ne l'autoriserons donc qu'en mode différé (remote batch, par exemple) de manière à limiter les coûts et l'occupation du terminal.

Ces états statistiques, à l'instar des précédents, sont parfois destinés à des personnes extérieures (Ministère, chercheurs, autres C.A.P., ...) et, par conséquent, doivent être éditées, ce qui n'était pas a priori le cas de l'interrogation directe au vidéo.

A ce niveau, il serait également intéressant de pouvoir présenter en clair ou sous forme graphique, les résultats obtenus de même que cataloguer, sous forme de macros, les questions usuelles (statistiques d'activité annuelles, mensuelles) ou des types de question paramétrables.

### §3. Le programme d'aide au diagnostic.

Le troisième sous-système identifié se distingue par son caractère moins impératif. C'est une proposition de conception d'un outil de raisonnement moins indispensable que le catalogue des produits ou que les dossiers d'appel, mais qui, nous l'avons montré, peut s'avérer être un outil de contrôle ou de première orientation (dans certains cas extrêmes) particulièrement apprécié.

Pour cette dernière raison, l'équipe informatique qui, voici plus de cinq ans, a travaillé au centre, a senti la nécessité, en accord avec les responsables, de développer pareille application. Cette équipe s'en est allée, l'ordinateur disponible a changé, les programmes sont devenus inadaptés et, par ce fait, quelque peu oubliés. De ce fait, plutôt que de nous acharner à convertir ces programmes écrits par d'autres, pour un autre matériel et dans une optique différente, nous avons

préférré repartir à zéro et tracer les lignes maîtresses d'une nouvelle application orientée vers le mode interactif, facilement et rapidement manipulable. Mais, avant d'y venir, peut-être le lecteur sera-t-il curieux de connaître l'ancien système ou, du moins, ses fondements théoriques.

Rappelons encore que, pour différentes raisons, l'équipe qui l'a mis en place s'est disséminée et que nous n'avons pu contacter aucun de ses membres.

### 3.1. L'ancien système.

Développé selon la plus simple méthode logique, il se fondait sur une matrice (150,800) synthétisant l'ensemble des connaissances. Pour chacun des 150 produits ou classes de produits recensés, le terme générique de cette matrice déterminait l'absence ou la présence d'un des 800 symptômes retenus.

Cette richesse et cette précision quant à la définition et l'identification des symptômes devait engendrer bien des confusions lorsqu'un médecin spécialiste du centre traitait avec une personne quelconque.

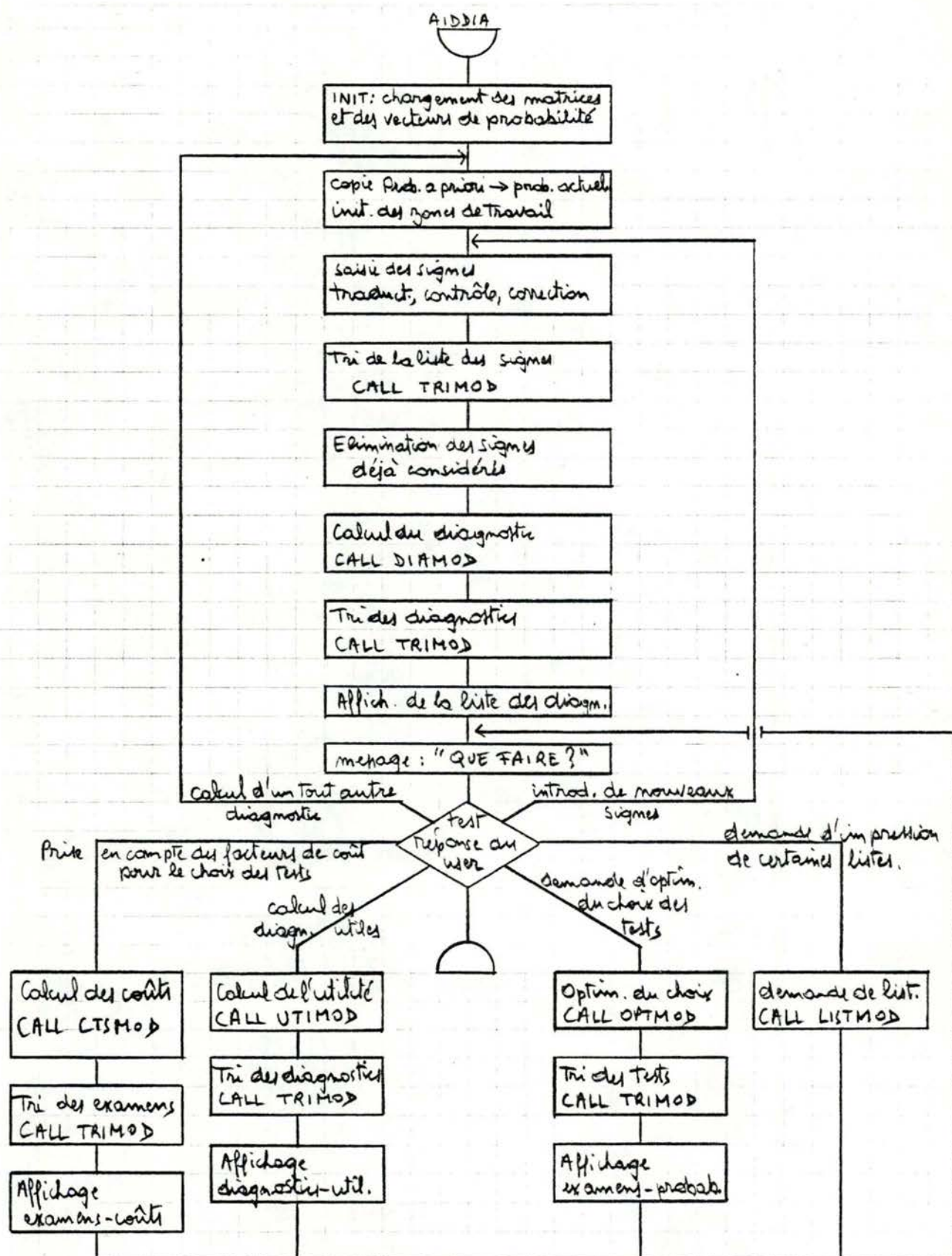
Un algorithme comparait ensuite le syndrome proposé par l'utilisateur (en mode différé !) à l'ensemble de la connaissance et en déduisait, après quelques ajustements, un ou plusieurs diagnostics (diagnostic = intoxication par absorption de tel produit). Aucun commentaire, aucune nuance n'émaillait cette liste de valeurs. Il appartenait donc au médecin d'interpréter ou d'introduire un syndrome plus riche ou plus élémentaire, selon son désir d'élaguer ou d'étoffer la liste obtenue.

### 3.2. Le système d'AIDe au DIagnostic proposé (AIDDIA).

Nous entrons directement dans le vif du sujet, en laissant au lecteur averti le soin d'analyser l'ordinogramme général (cfr. page suivante), manière plus que tout autre aisée de représentation d'une logique de raisonnement appliquée à un traitement.

Par la suite, nous fournirons quelques explications méthodologiques, ainsi qu'une série de remarques pratiques plus accessoires.





3.2.1. Routine TRIMOD.

- fonction : Trier un vecteur X de N éléments selon la valeur d'un indicatif A.
  - Si A = 0, tri en ordre croissant des valeurs absolues des éléments du vecteur X.
  - Si A = 1, tri en ordre décroissant des valeurs de X selon les valeurs des éléments correspondants du vecteur D (vecteur de N éléments).
- paramètres d'appel :
  - l'indicatif A,
  - le nombre entier N,
  - le vecteur X des éléments à trier,
  - le vecteur D, critère de tri lorsque A = 1.

3.2.2. Routine DIAMOD.

- fonction : Application de la formule de Bayes dans sa version :

$$P(D_j/S_1^+, S_2^-, \dots) = \frac{P^0(D_j) * P(S_1/D_j) * Q(S_2/D_j) * \dots}{\sum_k P^0(D_k) * P(S_1/D_k) * Q(S_2/D_k) * \dots}$$

(cfr. exposé théorique des méthodes probabilistes).

- paramètres d'appel :
  - la liste des nouveaux signes à incorporer dans le calcul du diagnostic,
  - les dernières probabilités calculées ( $P^0(D_j)$  pour cette itération du calcul),
  - la matrice des probabilités conditionnelles,
  - le nombre exact de nouveaux signes à prendre en compte.

3.2.3. Routine UTIMOD.

- fonction : Calculer une proposition de diagnostics basée sur une mesure de l'utilité. Pour chaque diagnostic  $D_i$ , il faut appliquer la formule :



$$Util(D_i) = \sum_j P(D_j) * U(i, j)$$

où l'élément  $U(i, j)$  de la matrice  $U$  représente l'utilité (inverse du risque) de porter le diagnostic  $D_i$  lorsque le diagnostic réel est  $D_j$ .

#### 3.2.4. Routine OPTMOD.

- fonction : Calcul de l'apport moyen d'information (par la fonction entropie) apporté par chaque test ou symptôme non encore pris en considération.  
(cfr. exposé théorique des méthodes de choix d'examen complémentaires).
- paramètres d'appel :
  - vecteur des probabilités des tests,
  - vecteur des probabilités actuelles des diagnostics,
  - vecteur des tests ou symptômes non encore pris en considération,
  - nombre de ces tests.

#### 3.2.5. Routine LISTMOD.

- fonction : Commander l'édition de certains résultats intermédiaires ou finaux :
  - les signes entrés,
  - les diagnostics probables,
  - les diagnostics utiles,
  - les tests probables,
  - les tests pondérés par les coûts.
- paramètres d'appel :  
Il peut s'agir de n'importe quels vecteurs de symptômes, de tests ou de diagnostics, de coûts, d'utilités ou de probabilités.

### 3.2.6. Routine CTSMOD.

- fonction : Fonction à paramétrer lors de l'initialisation du système ou lors de l'appel du programme, elle peut agréger, pour le choix des tests complémentaires, des facteurs aussi divers que :

- le coût financier du test,
- le risque encouru,
- le temps d'attente imposé par ce test,
- le gain d'information,
- ...etc.

A chacun de ces facteurs, la phase de paramétrage doit associer un coefficient de pondération pour les intégrer dans un simple modèle additif, ou déterminer une fonction de coût global plus élaborée.

Alors que le coût financier et le temps d'attente sont des données brutes pour chaque test, le gain d'information et le risque encouru dépendent des probabilités actuelles des diagnostics. Le premier de ces deux facteurs se calcule par la routine OPTMOD, tandis que le second le sera, à peu de choses près, par la routine UTIMOD.

- paramètres d'appel :

Outre les différents vecteurs des facteurs de coût et de la matrice associée au calcul des risques, nous devons citer :

- le nombre de tests concernés par cette évaluation des coûts,
- la liste de ces tests,
- les probabilités actuelles des diagnostics,
- les probabilités a priori des diagnostics,
- la matrice des probabilités conditionnelles,
- le vecteur des probabilités des tests.

### 3.3. Remarques.

Nous avons donc conçu un modèle assez général dont la caractéristique essentielle est sa possibilité de paramétrage.

- En dehors de la phase d'exécution proprement dite, l'uti-



lisateur peut introduire des facteurs de coût, modifier les matrices et vecteurs de probabilités, de risques, d'utilités, ainsi qu'exprimer des modèles de fonction de coût global différentes.

- Pendant l'exécution, il peut à tout moment orienter le déroulement du programme selon ses désirs. Il peut choisir une des fonctions de coût global qu'il aura préalablement cataloguées, ou baser le calcul du diagnostic sur l'hypothèse d'équiprobabilité des diagnostics.

L'utilisation de la méthode probabiliste et le rejet de la méthode logique préalablement implémentée se justifie ainsi :

- Nous déterminons, outre une liste de diagnostics, un critère hiérarchique représentatif des valeurs et des fréquences moyennes. S'il est démontré que la médecine n'est constituée que de cas particuliers, il nous semble cependant raisonnable et intéressant de comparer tout nouveau cas à une certaine moyenne et d'établir un classement des diagnostics possibles à partir de ces valeurs moyennes théoriques.

- L'estimation des probabilités a priori et des probabilités conditionnelles, précédemment présentée comme un désavantage certain, peut se réaliser aisément grâce à la fonction statistique implémentée dans le système de gestion des dossiers d'appel.

- Pour tenir compte de l'indépendance des symptômes, nous envisageons d'introduire un module spécifique dont la fonction serait d'éliminer tout signe défini comme non-suffisamment indépendant d'un autre signe déjà pris en compte dans le calcul du diagnostic. Pour ce faire, il nous suffit d'introduire, à l'intention de ce module, des listes de signes considérés comme dépendants entre eux.

exemple :

S1	0		
S2	S5	S14	0
S3	S6	0	
S4	0		
S5	S2	S14	0
S6	S3	0	
S.	..	...	..

Cette liste peut s'implémenter sous la forme de chaînes fermées, à condition que ces chaînes définissent une partition de l'ensemble des symptômes.

exemple : S1 S1  
S2 S5  
S3 S6  
S4 S4  
S5 S14  
S6 S3  
S. ..  
.. ..  
S14 S2



## CONCLUSION GENERALE.

Au terme de cette étude, nous voudrions, en guise de conclusion, placer deux remarques.

La première, pratique, nous permet de mettre en lumière l'aspect analytique de notre travail. Partant d'une situation réelle, nous avons conçu un système d'information susceptible de répondre adéquatement aux impératifs que nous avons recensés lors de nos enquêtes.

Nous avons ensuite procédé à l'analyse de ce système, analyse 'fonctionnelle' et 'organique', qui nous a vu aborder une bonne partie des problèmes qui y sont habituellement traités. Faute de disposer d'une méthode d'analyse des traitements qui nous convienne, nous nous sommes contentés de les décrire. Au fil de cette analyse, bien des aspects ont été éludés; nous pensons, entre autres, au planning des étapes de réalisation et de mise en place dans le milieu utilisateur, aux problèmes de rentabilité, ou plus généralement de coût. D'autres n'ont été que partiellement ébauchés, la définition d'un langage de manipulation orienté utilisateur, certains dictionnaires de codification.

Nous sommes conscients d'avoir aussi abordé certains problèmes relevant de l'analyse de programmation, trop tôt peut-être, mais notre analyse nous aurait semblé incomplète si nous n'avions pas, d'ores et déjà, porté notre choix sur des modèles d'implémentation.

Inéluctablement, les étapes de l'analyse, telles qu'elles sont usuellement définies, ne constituent pas des domaines disjoints; elles se caractérisent chacune par leurs spécificités, mais ne font pas moins l'objet d'une relative et certaine dépendance.

La seconde, théorique, nous permet, quant à elle, de lever un semblant de paradoxe. Le lecteur aura plus que vraisemblablement senti la confiance toute relative que nous accordions à ces méthodes d'aide au diagnostic. De ce fait, il

se sera étonné de nous voir proposer, à notre tour, l'implantation d'un outil que nous ne considérons qu'avec tant de précautions. Mais, ce même lecteur, aurait-il d'autres propositions, d'autres idées à faire valoir, lorsqu'il s'agit de diagnostiquer le cas fatal, le patient devant lequel le médecin est impuissant et inopérant ?

La méthode préconisée, au-delà des critiques qui lui sont adressées, comme toutes les autres d'ailleurs, offre l'avantage d'ébaucher des hypothèses, des solutions et en outre, ce qui la distingue des autres selon nous, présente, au travers des probabilités, un essai d'interprétation, à tester, qui peut s'avérer très utile, voire salvateur.

Cette justification introduit un problème d'ordre philosophique, sur lequel nous nous garderons de dissenter.

A la limite, la vie d'un homme, d'un seul être, ne vaut-elle pas le prix de cette initiative, de ce risque ?



# BIBLIOGRAPHIE.

- B1 - AFIFI A.A., SACKS S.T., LIU V.Y., WEIL M.H., SHUBIN H.  
 'Accumulative prognostic index for patients with  
 barbiturate glutemide and meprobamate intoxication.'  
The New England Journal of Medicine, vol 285, n°27  
 pp 1497-1502 (1971).
- B2 - ALPEROVITCH A., LELLOUCH J.  
 'L'aide au diagnostic. Bilan et perspectives.'  
Revue d'Informatique Médicale, vol 3, n°3  
 pp 205-215 (1972).
- B3 - ALPEROVITCH A., LELLOUCH J.  
 'Methods for aiding medical decision. Application to  
 Diagnosis of Round Intra-thoracic X-Rays Picture.'  
Computers and Biomedical Research, n°7 (1974).
- B4 - ALPEROVITCH A., LELLOUCH J.  
 'Value of computer diagnosis of a single thoracic  
 X-Ray circular opacity.'  
Biomedicine 1974, n°20.
- B5 - ALPEROVITCH A., LE MINOR M., LELLOUCH J.  
 'The choice of a strategy for the selection of com-  
 plementary investigations : a study by a simulation.'  
Medinfo 77, SHIRES/WOLF eds, IFIP North-Holland  
 Publishing Company (1977), pp 133-136.
- B6 - ASSELAIN B.  
 'Le concept d'utilité dans la décision médicale.  
A propos d'une expérience en neurologie.'  
 thèse (université Paris V) (1977).
- B7 - ASSICOT M., BOHUON C.  
 'Problems with on-line data acquisition systems in a  
 clinical chemistry laboratory.'  
Medical Informatics, vol 1, n°1, april 1976.
- B8 - BERLEUR J., POULLET Y.  
 'Informatique et vie privée. Essai d'une approche  
 comparative. Perspectives de solution.'

Journal des Tribunaux, 93-ème année, n°5028, fév. 78.

B9 - BERLEUR J.,POULLET Y.

'Informatique et vie privée : Enjeux réels.'

Interdoc, Bruxelles sept. 1978, n°1.

B10- BERNARD J.

'The medical decision.'

Decision making and medical care. Can information science help ? DE DOMBAL/GREMY eds, IFIP North-Holland Publishing Company (1976).

B11- BOUCKAERT A.

'Réflexions sur le diagnostic médical. Les campagnes de dépistage.'

Revue d'Informatique Médicale, vol 2, n°3 (1971)  
pp 243-251.

B12- BOUCKAERT A.

'Utilisation conjointe de l'algèbre booléenne et des méthodes statistiques pour le diagnostic médical.'

IRIA - Journées d'Informatique Médicale,  
1972 tome II.

B13- CARD W.I.

'The diagnosis process.'

Medical Computing. Progress and Problems.

ABRAMS M.E. (1969) Chatto and Windus for the  
British Computer Society.

B14- CARD W.I.,RUSINKIEWICZ M.,PHILLIPS C.I.

'Estimation of the utilities of states of health with different visual acuities using a wagering technique.'

Decision making and medical care. Can information science help ? DE DOMBAL/GREMY eds, IFIP North-Holland Publishing Company (1976).

B15- CARD W.I.,GOOD I.J.

'The estimation of implicit utilities of medical consultants.'

Math. Bioscience, vol 7,n°8, pp 45-54

B16- CROFT D.J.

'Is computerized diagnosis possible ?'



Computer. Biomed. Research, vol 5, 1972, pp 351-367

B17- DUCROT H., GREMY F.

'Le diagnostic assisté par ordinateur.'

La Revue du Praticien, 1973, tome 23, n°29,  
pp 2621-2633.

B18- DUSSERRE L., NAKACHE J-P., GREMY F.

'Tactical aspects in the use of linear discriminant analysis.'

Decision making and medical care. Can information science help ? DE DOMBAL/GREMY eds, IFIP North-Holland publishing Company (1976).

B19- FABRE R., TRUHAUT R.

'Précis de Toxicologie.'

tomes I et II, SEDES Paris

B20- FISKUS M.

'Une version nouvelle des programmes SINBAD.'

AFCET-IRIA - Journées Banque de Données 1971  
(Aix en Provence).

B21- FISKUS M.

'Bilan informatique d'une application de SINBAD II.'

Bulletin de Liaison du Club Banques de Données,  
n°15, july 1975, pp 138-168.

B22- FISKUS M., GERDAY R.

'The values of a dialogue for achieving optimum operation of scientific data banks.'

First International On-line Information Meeting,  
London (décembre 1977), One-Line Review.  
pp 173-178.

B23- FRAGU P., PATOIS E., HUBER C., LELLOUCH J.

'Original approach of the hyperthyroidism diagnosis with computer.'

Medinfo 74, North-Holland Publishing Company.

B24- FREJAVILLE J-P., BOURDON R.

'Toxicologie clinique et analytique.'

Flammarion.

B25- FREJAVILLE J-P., GOLDBERG M.

'Utilisation en mode conversationnel du catalogue d'un centre anti-poisons.'

Revue d'Informatique Médicale, vol 1, n°2, 1970  
pp 115-122

B26- GERDAY R.

'A system of direct access to information : the INSERM toxicological data bank.'

Medinfo 77, SHIRES/WOLF eds, IFIP North-Holland Publishing Company (1977), pp 307-311.

B27- GERDAY R.

'Une réponse aux besoins d'information en toxicologie. La banque de données de l'INSERM.'

4ème Colloque Informatique et Biosphère, Paris (1976)

B28- GERDAY R., BARBEROUSSE N.

'A data bank in toxicology using SINBAD programs : An economic solution to scientific information processing.'

1er Congrès de l'European Federation for Medical Informatics. Cambridge (septembre 1978).

B29- GERDAY R., FISKUS M.

'The evaluation of toxicological data. The methodological contribution of a data bank.'

Sixth Biennial International Codata Conference, Santa Flavia (Palermo), may 1978.

B30- GERDAY R., PINHAS N., RAVIART M.

'Une banque de données en toxicologie expérimentale peut-elle faire progresser les techniques d'évaluation des risques.'

Revue Int. Océanogr. Méd., tome L, 1978, pp 83-91.

B31- GINSBERG A.S.

'The diagnostic process viewed as a decision problem.'

Computer Diagnosis and Diagnostic Methods.

JACQUEZ ed, C. C. Thomas Publisher Springfield Illinois.

B32- GOLDBERG M., DOYON B.

'SARI : A user-oriented data-bank system for medical applications.'

Méth. Inform. Méd., vol 15, n°2, 1976.



- B33- GOLDBERG M., EMERY J.-P., GRAISELY B., DUCROT H., SANTINI C., VEZINET B., DOYON B., PLOTKINE M., PIGNARD P., FREJAVILLE J.-P.  
'Problèmes liés à la nature des informations et aux modalités d'exploitation dans la réalisation d'une banque de données sur les médicaments.'  
Colloques de l'IRIA. Informatique Médicale, Toulouse,  
tome II, mars 74, pp 287-300.
- B34- GREMY F., GOLDBERG M.  
'Decision making methods in medicine.'  
Informatics and Medicine 3, An advanced course,  
pp 419-459.
- B35- HABBEEMA J.D.F., HILDEN J., BJERREGAARD  
'The measurement of performance in probabilistic diagnosis.'  
Methods of Information in Medicine, vol 17, n°4 (oct.78).
- B36- JACQUEZ J.A., NORUSIS M.J.  
'The importance of symptom non-independance in diagnosis.'  
Decision making and medical care. Can information science help ? DE DOMBAL/GREMY eds, IFIP North-Holland Publishing Company (1976), pp 379-392.
- B37- KOHLER C.O., WAGNER G., WOLBER U.  
'Interactive data processing in medicine. Man-machine dialogue. (Bibliography).'  
Methods of Information in Medicine, vol 15, n°2,  
1976, pp 102-120.
- B38- LABIN E.  
'Les banques de données dans le domaine scientifique et technique.'  
BNIST (Paris, La Documentation Française) 1976.
- B39- LANGE H.J.  
'Decision making by doctors and aided by computers.'  
Decision making and medical care. Can information science help ? DE DOMBAL/GREMY eds, IFIP North-Holland Publishing Company (1976), pp 11-22.
- B40- LAUWERYS R.  
'Précis de toxicologie industrielle et des intoxications professionnelles.' Duculot.

B41- LUSTED L.B.

'Clinical decision making.'

Decision making and medical care. Can information science help ? DE DOMBAL/GREMY eds, IFIP North-Holland Publishing Company (1976), pp 77-100.

B42- LELLOUCH J.,ALPEROVITCH A.

'Une note sur l'application de l'algorithme de FORD à un problème de diagnostic.'

Revue Française d'Informatique et de Recherche Opérationnelle, vol 5, n°1, 1971, pp 57-59.

B43-MARTINEZ ALONSO J.R.,MANRIQUE LEGAZ A.,ORTIZ BERROCAL J.

'Automatic diagnosis with computer.'

Colloques de l'IRIA. Informatique Médicale, Toulouse, tome 2, mars 73, pp 35-60.

B44- MATHIOT M.

'Réactions de l'entourage à l'implantation de systèmes informatiques.'

Colloques de l'IRIA. Informatique Médicale, Toulouse, tome 2, mars 75, pp 117-127.

B45- MATTEI M..

'Les ordinateurs à l'aide du diagnostic médical. Application à la toxicologie.'

thèse présentée à la faculté mixte de Médecine et de Pharmacie de Grenoble. (1969)

B46- MATTEI M.,FAURE J.,YACOUB M.

'Aide de l'informatique au diagnostic médical. Application à la toxicologie clinique d'urgence.' (1971)

Revue d'Informatique Médicale, vol 2, n°2, pp 139-147.

B47- MATTEI M.,VALOIS J.,DODELIER C.,DUQUESNEL P.,HERNICOT P.

'Implantation du système Socrate dans un service de toxicologie clinique.'

Colloques de l'IRIA. Informatique Médicale, tome 1, 1972.

B48- NAKACHE J-P.

'What is data analysis and its current medical applications ?'

Medcomp 77.



- B49- NAKACHE J-P., LORENTE P., BENZECRIJ-P., CHASTANG J-F.  
 'Aspects pronostiques et thérapeutiques de l'infarctus myocardique aigu compliqué d'une défaillance sévère de la pompe cardiaque. Application des méthodes de discrimination.'  
Les Cahiers de l'Analyse de Données, vol 2, n°4,  
 1977, pp 415-434.
- B50- RECTOR A.L., ACKERMAN E.  
 'Rules for sequential diagnosis.'  
Computer. Biomed. Research, vol 8, 1975, pp 143-155.
- B51- SALAMON R.  
 'Aide à la décision et enseignement médical à propos d'une expérience en neurologie.'  
 thèse de doctorat d'Etat en biologie humaine,  
 (université Paris VI) 1976..
- B52- SALAMON R., BERNADET M., GREMY F..  
 'Ternary algebra : theoretical aspects and possible applications.'  
Decision making and medical care. Can information science help ? DE DOMBAL/GREMY eds, IFIP North-Holland Publishing Company (1976), pp 505-514.
- B53- SALAMON R., DEROUESNE C., SAMSON M., BERNADET M., GREMY F.  
 'Decision-making aids used to determine the content of medical teaching.'  
id.
- B54- SALAMON R., GREMY F., BERNADET M., THONNET M., DEROUESNE C., SAMSON M.  
 'Une expérience d'aide au diagnostic en neurologie avec un objectif pédagogique.'  
Medical Data Processing Symposium, Toulouse  
 LAUDET/ANDERSON/BEGON eds, pp 549-559.
- B55- SALAMON R., MIHOUT B., SAMSON M., BERNADET M., GREMY F.  
 'Aide à la décision en neurologie : Application au diagnostic topographique.'  
Medical Computing, LAUDET/ANDERSON/BEGON eds  
 TAYLOR AND FRANCIS - LONDON , pp 391-398.
- B56- SIMONIN R., ROUX M., GOUVERNET J.

'Initiation à l'aide au diagnostic.' (1973)

Revue d'Informatique Médicale, vol 4, n°1, pp 45-54.

B57- SUNSHINE I.,GOVAERTS A.,GAULTIER M.,ROCHE L.,VINCENT V.

'Les centres anti-poisons dans le monde.'

Collection de Médecine légale et de Toxicologie  
médicale, Masson et Cie editeurs.

B58- TAYLOR T.R.

'Clinical decision analysis.'

Meth. Inform. Med., vol 15, n°4, 1976, pp 216-224.

B59- VAN BRUNT E.

'Factors which limit or constraint the development of  
medical data processing.'

Medical Informatics, vol 1, n°4, déc. 1976.

B60- VINCENT V.,MICHELA G.,DUCLOUX B.

'Une utilisation de l'ordinateur en toxicologie.

L'expérience des centres anti-poisons.'

Medical Data Processing, LAUDET/ANDERSON/BEGON eds,  
TAYLOR AND FRANCIS - LONDON , pp 157-167.

B61- ZERAFFA J.

'L'INSERM et l'information biomédicale.'

Bulletin des Bibliothèques de France, n°2, 1972,  
pp 35-66.



ANNEXES.

CODE	VARIABLES	classe 1	classe 2	classe 3	classe 4	classe 5
RV	SEXE	masculin	féminin			
AC	AGE (années)	57	50-68	70		
PR	INFARCTUS DANS LES ANTECEDENTS	absence	présence			
TS	DATE DE LA PREMIERE OBSERVATION PAR RAPPORT AU DEBUT DE L'INFARCTUS	1 jour	2 jours	3 jours		
A/I	LOCALISATION DE L'INFARCTUS	anterior	postérieur			
HR	FREQUENCE CARDIAQUE (Systoles / min)	82	84-96	98		
CI	INDEX CARDIAQUE (l/min.m <sup>-2</sup> )	1.23	1.24 - 1.58	1.58 - 1.82	1.83 - 2.40	2.41
SI	INDEX SYSTOLIQUE (ml/systole.m <sup>-2</sup> )	13.8	13.8 - 17.4	17.5 - 21.7	21.8 - 26.8	27.2
DA	PRESSION DIASTOLIQUE AORTIQUE (mmHg)	58	58 - 67	68		
PA	PRESSION AORTIQUE MOYENNE (mmHg)	88	70-81	82		
R	RESISTANCE ARTERIELLE SYSTEMIQUE (dyn.sac.cm <sup>-5</sup> .m <sup>-2</sup> )	2156	2172-2808	2846-3164	3173-3897	4031
LP	PRESSION DE REMPLISSAGE VG (mmHg)	14	15-17	18-20	21-24	25
PP	PRESSION ARTERIELLE PULMONAIRE MOYENNE (mmHg)	18	20-23	24-27	28-31	32
RP	PRESSION TEDIASTOLIQUE VD (mmHg)	7	8-11	12		
PR	RESISTANCE PULMONAIRE TOTALE (dyn.sac.cm <sup>-5</sup> .m <sup>-2</sup> )	787	807-1108	1131-1608	1610	
LS	TRAVAIL GAUCHE SYSTOLIQUE (g.m/m <sup>2</sup> )	10.48	10.72-14.83	15.28-24.0	24.70	
MS	TRAVAIL DROIT SYSTOLIQUE (g.m/m <sup>2</sup> )	2.74	2.82-4.15	4.16-6.15	6.22	
L*	PUISSANCE GAUCHE MINUTE (g.m/min.m <sup>-2</sup> )	1022	1031-1325	1328-2088	2131	
PM	PUISSANCE DROITE MINUTE (g.m/min.m <sup>-2</sup> )	242	246-387	388-564	573	
L1	INDEX GAUCHE (g.m/m <sup>2</sup> .mmg <sup>-1</sup> )	0.44	0.45-0.64	0.67-0.83	0.85-1.85	1.72
PI	INDEX DROIT (g.m/m <sup>2</sup> .mmg <sup>-1</sup> )	0.20	0.21-0.35	0.36-0.53	0.54-1.14	1.16
S/D	SURVIE/DÉCÈS	survivant	décédé			
CD	TRUPLES DE CONDUCTION	absence	présence			
RV	INDEX DE GENERATION DE PRESSION VENTRICULAIRE DROITE (sans dimension)	0.94	1.05-1.82	1.83-2.88	3.00	

Tableau 1 : Liste des variables de l'étude, du code utilisé dans l'AFC et du découpage en classes



Annexe 2.

(Les informations marquées d'un astérisque\* sont utilisables pour la sélection et l'édition, les autres pour l'édition seulement).

1 - Fichier "Principes actifs"

- Terminologie : dénomination principale, autres dénominations, dénominations génériques, dénomination commune française, dénomination internationale proposée, dénomination internationale recommandée, sels et dérivés.
- Caractéristiques chimiques : structure chimique, classe (s) chimique (s)\*
- Caractéristiques pharmacodynamiques : classe (s) pharmacodynamique (s)\*
- Mécanisme d'action, pharmacocinétique (absorption, répartition, métabolisme, élimination). Action perturbatrice sur les examens de laboratoires \*
- Thérapeutique : effets recherchés \*, indications thérapeutiques, effets secondaires (nature \* fréquence \* conditions d'apparition exclusive, \* conditions favorisantes), risque de pharmacodépendance \*, contre indications (nature \*, en cas d'administration par voie : \*) précautions d'emploi (nature \*, en cas d'administration par voie : \*), interactions médicamenteuses (effet thérapeutique accru par association avec : \*, effet thérapeutique diminué par association avec : \*, associations conseillées (nature, en cas de :) associations à déconseiller (nature, en cas de :))
- Toxicologie : principaux symptômes\*, traitement d'urgence de l'intoxication aiguë
- Maniement usuel : posologie, mode d'emploi, législation, \*
- Noms des spécialités contenant la substance seule (Marché commun, Suisse, U.S.A) ou en association (France)
- Références bibliographiques, date de dernière mise à jour \*.

## 2 - Fichier Spécialités

- Terminologie : nom de la spécialité, forme pharmaceutique \*, numéro d'expérimentation ; anciennes dénominations, nom générique, nom et adresse du fabricant \*
- Composition : principes actifs, excipients, \* conservateurs \*
- Conservation : durée \*, précautions particulières \*
- Présentation : vente public (contenance, régime de délivrance \*, prix public, nature du récipient \*), présentation pour hôpitaux (contenance, régime de délivrance \*, prix, nature du récipient \*)
- Thérapeutique : propriétés thérapeutiques \*, effets biopharmaceutiques \*, indications thérapeutiques, effets secondaires (nature \*, fréquence \*, conditions d'apparition exclusive \*, conditions favorisantes \*) Précautions d'emploi \*, contre indications \*, associations conseillées (nature \*, en cas de : \*) associations à déconseiller (nature \*, en cas de : \*)
- Maniement usuel : voies d'administration \*, posologie (posologie habituelle, posologie maximum par prise et par jour, posologies particulières), mode d'emploi, incompatibilités physicochimiques
- Réglementation \* : brevets, numéro de visa, date de mise sur le marché, réglementation des prix \*, collectivités \* expertises
- Date de dernière mise à jour \*



ERGOTAMINE

Vasoconstricteur - Alcaloïde de l'ergot

Médicament

Sel utilisé : Tartrate.

Noms commerciaux : Gynergène, Fémergin, Cafergot (p.p.), Medihaler ergotamine.

Caractéristiques physiques : Poudre blanche ou cristaux incolores, faiblement soluble dans l'eau et l'alcool (1/500).

pratiquement insol. dans l'éther et le chloroforme. F° : 203°C. (déc.)

Pour obtenir une solution aqueuse limpide, il est nécessaire d'ajouter une trace d'acide tartrique.

Sensible à la lumière, la chaleur, l'oxydation.

Usages : Traitement de la crise migraineuse; au cours du post-partum (hémorragies, subinvolution utérine après expulsion du placenta).

Doses usuelles : selon (1) dans le traitement de la migraine : p.o. : 4 à 5 mg à la fois, puis 2 mg/h.

Ne pas dépasser 11 mg par crise.

Sublingual : 3 à 4 mg à la fois puis 1 à 2 mg/h. Ne pas dépasser 10 mg par crise.

selon (2) : 2 à 3 mg/j. p.o.

selon (9) : ne pas dépasser 6 mg/j. ou 12 mg/semaine.

I.M. ou S.C. : 0,25 à 0,5 mg/j. Max. : 1 mg/j.

Voie rectale : 2 mg à répéter 1 à 2 fois par crise. Ne pas dépasser 6 fois 2 mg par semaine.

Aérosol : 0,36 mg par inhalation, à répéter toutes les 5 minutes. Max. : 6/j. et 15/semaine.

D.M. : 1 mg par prise, 2 mg/j. (p.p. IV).

Contre-indications : méthysergide, nicotine, sympathicomimétiques, grossesse, affection hépatique ou rénale grave, artérites, infections, hyperthyroïdie.

Mécanismes physiopathologiques : Bloque les récepteurs  $\alpha$ -adrénergiques (à doses non utilisables en clinique).

Actions vasoconstrictive propre et toxique sur les endothelia vasculaires. Action utérotonique prolongée.

Activité p.o. discutée (1)(2) : résorption lente, incomplète et inconstante.

Bien résorbé par les muqueuses (buccale et rectale). Action relativement lente (même en I.V.).

Stimule, et à forte dose paralyse, les terminaisons motrices sympathiques (vasculaires, intestinales, utérines).

Toxicité : D.L. 50 (mg/kg) | I.V. | S.C.

	I.V.	S.C.
rat	62-80	100-150
lapin	3,6	
souris	52	

Chez l'homme : décès chez un adulte avec 26 mg en plusieurs jours; de même après 0,5 à 1,5 mg I.V.;

avec 12 mg chez un enfant de 14 mois; début de gangrène avec 40 mg en 5 jours; gangrène et amputation

des jambes avec 7 mg à raison de 3 mg/j. p.o. (6).

Effets secondaires : (rares aux doses thérapeutiques)

digestifs: nausées, vomissements,

neuro-psychiques : convulsions,

généraux : oedème et prurit chez de rares sujets sensibilisés; céphalées (qui poussent à augmenter les doses),

mictions impérieuses, vasodilatation. Douleurs musculaires et impression de faiblesse dans les jambes;

picotements dans les doigts et les orteils; crises d'angor (8); tachy- ou bradycardie; hypercoagulabilité;

Références :

1) Goodman et Gilman, 882.

2) Monographie Sandoz.

3) Jones E.M., Williams B; Brit. Med. J., 1966, I, 466.

4) Dreisbach, V, 274.

5) Carter, E.R.; Brit. Med. J., 1958, 11, 1452.

6) Cameron E.A. et al., Brit. Med. J., 1960, 11, 28.

7) Semi-annual Tabulation on blood dyscrasias, Am. Med. Assoc., 1962, 6, n°2.

8) Goldfischer, New Engl. J. Med., 1960, 282, 860.

9) Martindale, 590.

Tous droits réservés pour tout pays.

Centre belge anti-Poisons Août 1967.

Annexe 4.L'Expérience de la Banque Européenne des  
Centres Anti-Poisons.A. Origine.

Cette réalisation est née d'un besoin réel qui s'est fait sentir dès la fin des années soixante. L'IMA (Information Médicale Automatisée), centre de documentation de l'INSERM (Institut National de la Santé et de la Recherche Médicale) venait de constituer et d'expérimenter une banque de données en Enzymologie (1968-70), gérée par un logiciel spécifique, les programmes SINBAD (Système Informatique pour Banque de Données).

Parmi les autres disciplines possibles d'application figurait en assez large priorité la toxicologie, alors en pleine croissance.

L'accroissement considérable du nombre de produits posait des problèmes de plus en plus complexes de maîtrise et d'accès rapide à pareil volume d'informations, dans des secteurs aussi variés que la Santé Publique, l'agriculture ou l'industrie. Fait déterminant, les utilisateurs potentiels ne manquaient pas.

Une étude des conditions économiques de fonctionnement de la Banque de Données démontra l'utilité d'une subdivision du domaine toxicologique en cinq secteurs caractérisés par leurs besoins, leurs documents :

- les centres anti-poisons,
- les services hospitaliers de réanimation et de désintoxication,
- les laboratoires de toxicologie expérimentale,
- les laboratoires de recherche sur les mécanismes d'action,
- les services de Santé Publique étudiant les effets à long terme.

Priorité fut accordée aux centres anti-poisons.



## B. Source des données.

A cette époque, divers centres européens entretenaient des relations constantes. Leur désir de coopération et leur expérience individuelle constituaient de sérieux atouts. Douze centres participèrent au programme : les centres français de Paris, Lyon, Marseille, Lille, Nancy, Rennes, Strasbourg, Toulouse, le centre de Bruxelles, celui de Zurich, les centres allemands représentés par Mayence et les italiens, par Rome.

La première étape du travail fut la standardisation d'une fiche de recueil des appels téléphoniques, enregistrement de tous les renseignements jugés indispensables, ou du moins utiles, pour une exploitation ultérieure. Ces fiches, remplies lors de chaque demande adressée à l'un des centres, et la littérature constituaient les deux principales sources de données.

Un exemplaire de la double fiche d'appel figure en annexe 7.

## C. Saisie des données.

La seconde étape du travail fut le recueil effectif de ces dossiers par chaque centre et l'envoi du double (rendu anonyme) de ces fiches à l'IMA (1972). La structuration des données exigée par le logiciel SINBAD (cfr annexe 8) fut dès lors opérée, suivie de l'enregistrement proprement dit. En deux ans, l'IMA recut près de 30.000 dossiers comportant, en moyenne, 31 données.

Dès 1974, le centre put procéder à l'interrogation de la banque de manière satisfaisante.

## D. L'outil informatique.

L'outil employé est le logiciel SINBAD, créé et mis au point par M.FISKUS (B20, B21) avec l'aide de I.OSOROVITZ, amélioré en une version SINBAD II plus performante.

Nous exposerons les principes généraux de ce système de gestion dans le chapitre concernant les banques de données toxi-

cologiques de l'INSERM, tout en précisant que les quelques améliorations apportées visent l'interrogation de la banque, sans modifier la structuration des données.

### E. Exploitation.

Dans le respect des impératifs posés, l'interrogation de cette banque s'exécute en mode conversationnel et présente effectivement un accès non privilégié aux données. Problème de compromis classique, l'interrogation et les procédés d'économie d'écriture et de codage très avantageux empêchent la mise à jour instantanée, qui n'était pas, précisons-le, un objectif essentiel.

#### §1. Interrogation.

- Procédure simple, rapidement assimilable par un non-spécialiste (malgré la formulation entièrement codée).
- Accès direct.
- Accès non privilégié (pas de clés d'accès spécifiques), le nombre de critères de sélection n'est pas limité.
- Temps de réponse pratiquement indépendant de la complexité de la question. Les transferts de données en mémoire sont limités au minimum.
- Accès soit à des données précises, soit à des combinaisons de données, grâce à l'usage d'opérateurs booléens (et/ou/sauf), arithmétiques (<, >, %-compris entre), synthétiques (M-identité/D-différentiation/N-comptage).

#### §2. Mise à jour.

Le système SINBAD rend impossible les mises à jour directes, car la compression maximum effectuée sur le BLOC-LIGNES (l'un des principaux fichiers de la banque), l'enregistrement en longueurs variables et l'éclatement des données vont précisément à l'encontre des techniques sous-jacentes à la mise à jour instantanée, à savoir l'insertion de places libres et les multiples pointeurs. Il s'agit donc de regrouper les mises à jour en sessions,



lors desquelles le système doit procéder à des recopies importantes pour des insertions parfois minimes. La fonction de mise à jour se limite à l'introduction de nouveaux éléments.

#### F. Utilisation.

L'exploitation en coopération offrait bien des avantages :

- partage des coûts de recherche, de développement et de mise en place du système d'exploitation (éventuellement de l'investissement de base),
- accès à un plus grand nombre de données de fiabilité plus grande car basée sur davantage d'observations,
- meilleure organisation des différents centres grâce à de bien meilleures études statistiques et épidémiologiques,
- standardisation de la terminologie utilisée.

A cette époque, la mise en place d'un réseau de terminaux intelligents posait cependant quelques inconvénients que l'évolution technologique a résolu depuis lors.

Il ne restait en fait alors qu'à définir la structure juridique d'une telle association internationale de service public. Hélas, des intérêts divers ont prévalu à la réussite de cette expérience. L'esprit de coopération a fait défaut à certains échelons, chacun défendant son optique de partage et de propriété des données. L'expérience de la banque européenne des Centres anti-poisons en est resté au stade de l'expérimentation.

#### G. Rentabilité. Résultats.

La fin prématurée de l'expérience ne met pas en cause le système informatique tel qu'il a été conçu et appliqué ; elle n'en a pas donné le temps.

Mais un bilan peut néanmoins être dressé au niveau de ce que chaque centre en a tiré :

- pour les uns, un premier aperçu et donc une base de lancement pour la mise en place d'un système individuel,
- pour d'autres, une comparaison et une mise au point de leurs méthodes pratiques,
- pour d'autres encore, une réorganisation plus rationnelle et plus efficace de leur système d'information grâce notamment à l'implantation du système des fiches d'appel complètes et exploitables dans un avenir plus ou moins éloigné.



## Les Banques de Données Toxicologiques de l'INSERM.

### A. Origine.

Ainsi que nous l'avons déjà vu, le développement extraordinaire de la toxicologie, parallèlement à d'autres domaines médicaux, a placé les spécialistes devant l'impossibilité de maîtriser une telle masse de données.

Certes, il existait alors des systèmes de documentation automatisée, tel MEDLARS (Medical Literature Analysis and Retrieval System), rebaptisé MEDLINE (Medlars 'on-line'), susceptibles de répondre à des combinaisons de mots-clés définis par l'utilisateur, par une liste assez complète de références bibliographiques pertinentes et de leur thème. Mais la démarche du chercheur, un peu simplifiée certes, ne s'arrêtait pas là. Il lui fallait trouver les articles mentionnés, puis analyser les documents pour en finalement tirer les renseignements qu'il désirait. Dans certains cas, pareille procédure finissait par l'exaspérer, lui qui préférait, pour un gain d'intérêt et de temps, procéder lui-même à une expérience, sans exploiter celles des autres. Il était donc urgent de concevoir des systèmes qui puissent fournir l'information contenue dans les articles en un minimum de temps.

La banque des données était l'outil rêvé.

Une enquête entreprise auprès d'institutions de recherche (600 questionnaires distribués), de firmes pharmaceutiques (400) et de l'industrie chimique (1600), suivie d'une vingtaine d'interviews semi-directifs puis de sessions d'études rassemblant quelques experts, révéla deux types d'utilisateurs potentiels, caractérisés par des besoins divergents : - les milieux administratifs et industriels : information simple et validée résumant l'état des connaissances, accompagnés de la source bibliographique,

- les laboratoires de recherche : information très précise et très détaillée ou synthèse de cette connaissance.

Ces deux catégories d'utilisateurs ont déterminé la création de deux banques indépendantes :

- la Banque de Données Synthétisées (B.D.S.)
- la Banque de Données Détaillées (B.D.D.) ou Banque de Données en Toxicologie expérimentale.

Deux impératifs furent associés à ces besoins. Ce sont, d'une part l'accès direct permettant la réponse immédiate et l'exploration des fichiers par un authentique dialogue homme-machine, et de l'autre, l'accès isotrope, ce qui revient à dire qu'aucune classe de données n'est privilégiée au niveau de l'implémentation.

## B. Source des données.

### §1. B.D.S.

Toutes les données proviennent de monographies, d'ouvrages ou de normes internationales. Les données rassemblées dans cette banque peuvent être regroupées en trois catégories :

- des données toxicologiques générales  
(toxicité (DL 50),...etc.),
- des données physico-chimiques  
(solubilité, densité, point d'ébullition,...etc.),
- des données officielles  
(concentration maximum admissible, normes d'emploi ou d'hygiène,...etc.).

### §2. B.D.D.

Elle rassemble les résultats d'expériences sur animaux, tirés essentiellement des comptes-rendus de laboratoire.

Cette banque contient des informations sur :

- le produit : nom, degré de pureté, origine,...etc,
- le document d'origine,
- les composants du produit,
- le sujet : espèce, souche, âge, poids, conditions de vie,
- les conditions d'administration du produit : mode



- d'administration, quantité, moment, rythme d'administration,...etc,
- le détail des résultats.

### C. Saisie des données.

Dans la plupart des applications informatiques, la saisie des données et leur contrôle constituent l'un des postes les plus délicats et les plus coûteux en temps et en argent. Il ne faut pas que la masse des données nuise à leur précision et à leur fiabilité. Cette remarque se révèle d'autant plus pertinente dans le cas de la Banque de Données Détaillées où l'objectif majeur exige un haut degré de finesse entraînant une organisation rigoureuse des données à l'entrée.

Pour répondre à cet impératif d'exactitude, l'analyse en profondeur des comptes-rendus d'expérience est confié à des scientifiques de haut niveau. Mais le moyen qui garantirait la meilleure information et qui limiterait considérablement les coûts de sa saisie consiste en l'envoi direct des formulaires d'analyse à l'auteur de l'article ou au laboratoire. D'autres moyens pourraient être le partage de la saisie entre divers organismes spécialisés ou la limitation aux produits les plus demandés.

Nous présentons en annexe 9, la structuration des données relatives à chacune des deux Banques.

### D. L'outil informatique.

#### §1. Hardware.

- ordinateur UNIVAC série 1100 de la faculté des Sciences de Paris installé à ORSAY
- un terminal unique est installé dans les locaux de l'équipe des Banques de Données de l'INSERM (Hôpital de Bicêtre, KREMLIN-BICETRE).

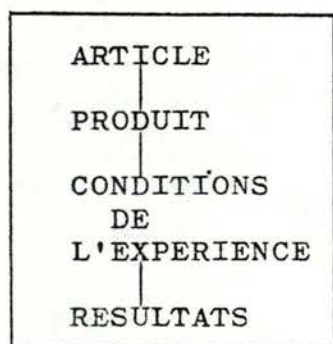
## §2. Software.

C'est l'ensemble des programmes SINBAD II qui gère les deux banques toxicologiques, de même que, précédemment, la banque des centres anti-poisons.

Quatre hommes-années ont suffi à la création et à la mise au point de ce logiciel spécifique, sans compter les modifications mineures qui y sont apportées périodiquement. Comme nous le verrons tout au long de la description que nous allons en faire, il s'agit d'un système optimisant, dans les limites assez contraignantes qu'il s'est fixées.

### 2.1. Structure fonctionnelle.

Nous pourrions schématiser la structure adoptée en une hiérarchie à 4 niveaux.



Cependant, SINBAD adopte une structure matricielle, réduisant les deux premiers niveaux en un référentiel double (un article, un produit), le troisième niveau déterminant, quant à lui, le numéro d'essai.

### 2.2. Groupes de classe.

Comme dans tout système de gestion de base de données, le regroupement de différentes classes de données offre certains avantages au niveau du traitement. Ces regroupements constituent les grilles des différentes banques de données. L'articulation de l'ensemble des données en grilles répond



aux relations sémantiques existant entre les différentes classes de données (on regroupe les classes qui ont le plus de chance d'être remplies simultanément).

### 2.3. Types de classes de données.

Cette distinction est, à notre connaissance, propre à SINBAD. Nous en démontrerons l'utilité plus tard.

- Une classe d'individus est une classe dont les valeurs possibles sont en nombre illimité.
- Une classe de caractéristiques comporte un nombre restreint de valeurs, en fait moins de 63.
- Une classe mixte est à l'origine une classe de caractéristiques dont le nombre de valeurs s'est, au fil du temps, révélé supérieur à 63. Cette classe offrant peu d'intérêt, nous n'en parlerons pas davantage.

### 2.4. Structure des fichiers.

La définition du type d'une classe de données détermine le fichier où elles seront enregistrées.

- L'article Y décrivant le produit ASPIRINE comporte 3 essais. Pour chacun de ces essais, nous disposons de certaines informations réparties dans les grilles. Toutes les grilles peuvent toujours être mises en relation par le double référentiel et l'essai, puisque ces renseignements sont repris dans chaque grille.
- L'ensemble des valeurs des caractéristiques, pour chaque grille, s'enregistre dans le BLOC-LIGNES qui est un fichier unique. Ce fichier ne contient en fait que des lignes de valeurs totalement indépendantes de l'essai et de la grille auxquels elles appartiennent.
- Un second fichier, le BLOC-IMAGE établit ce lien d'appartenance des différentes lignes de valeurs du BLOC-LIGNES au(x) numéro(s) de grilles dont elles relèvent.
- Le dernier lien à effectuer, le rattachement entre chacune

de ces paires n° grille-ligne de valeurs du BLOC-LIGNES et l'article-produit-essai est conservé dans les fichiers REFCOD-n, n représentant le numéro de la grille en question. Ainsi donc, un fichier REFCOD-i rattache certaines lignes du BLOC-LIGNES (celles qui relèvent de la grille i), aux différents articles-produits-essais dont elles dépendent.

- Les valeurs d'individus sont enregistrées de façon classique dans le BIBLIOCLAIR. Par article, un pointeur par grille détermine une suite de (n°s ligne du BLOC-LIGNES, n° segment du REFCOD-i qui identifie un article-produit-essai, les valeurs d'individus), le tout en longueur variable. Un ensemble article-produit-essai constitue un segment.

- Pour répondre aux impératifs de l'accès isotrope, un fichier inversé est constitué pour chaque valeur k prise par l'un des individus. Il s'agit des REFCLAIR-k. Dans chacun de ces fichiers inversés, un numéro de grille associé à un type d'individu pointe sur des segments constitués de (numéro de ligne du BLOC-LIGNES-numéro de segment d'un REFCOD-i).

## 2.5. Méthode d'économie d'écriture et de codage.

### 2.5.1. Codage des informations.

Les individus sont codés par une valeur numérique qui leur est attribuée selon leur moment d'apparition dans les énoncés. Les caractéristiques sont, quant à elles, codées selon leur rang de fréquence dans la classe à laquelle chacune d'elle appartient. Celle de fréquence la plus élevée reçoit le code de valeur de représentation binaire la moins élevée. Nous en verrons tout de suite l'utilité.

### 2.5.2. Formes d'économie.

Ces économies sont appliquées au niveau du BLOC-LIGNES, c'est-à-dire au niveau de l'enregistrement des lignes de valeur (énoncés) de caractéristiques de chaque grille.



### 2.1. L'économie d'attribution.

Une ligne identique pour plusieurs énoncés appartenant au même groupe de classes n'est écrite qu'une seule fois dans le BLOC-LIGNES. Les liens avec l'article-produit-essai (ou les) et avec la grille i concernée apparaissent respectivement dans le fichier REFCOD-i et dans le BLOC-IMAGE.

### 2.2. L'économie de superposition.

Deux lignes issues de groupes de classes différentes, éventuellement de longueur différente, exprimées par des suites de codes identiques, n'apparaissent qu'une fois dans le BLOC-LIGNES.

Les grilles dont elles proviennent peuvent se déduire des bits du BLOC-IMAGE indiquant les "propriétaires" de chaque ligne du BLOC-LIGNES.

### 2.3. L'économie de longueur.

Toute ligne du BLOC-LIGNES est caractérisée par une valeur de VPE (code de Valeur la Plus Elevée dans la ligne).

Le BLOC-LIGNES est subdivisé en 6 cadrans regroupant les lignes de VPE 63 à 32 pour le premier cadran, de VPE 31 à 16 pour le second,...etc.

Au niveau de l'accès, si une question porte sur une caractéristique de code 58, l'algorithme de recherche limitera ses investigations au premier cadran du BLOC-LIGNES.

L'économie de longueur consiste à coder sur 6 bits toutes les valeurs des lignes du premier cadran, sur 5 bits celles du second cadran,...etc.

## 2.6. Enchaînement des principaux modules.

- Un module de diagnostic analyse la question (syntaxe).
- Un module TRADON traite les éléments donnés.
- Un module EXCOD traite le fichier REFCOD en fonction des numéros de lignes fournis par TRADON.
- Un module EXINC extrait tous les éléments recherchés.
- Un module EXDINF répond à l'emploi des opérateurs.
- Un module de décodage.
- Un module d'édition (gestion des 1/0).

## E. Exploitation.

Nous ne reprendrons pas ici tout ce qui a été développé au chapitre précédent.

## F. Utilisation.

Au contraire de la banque destinée aux centres anti-poisons, les deux banques toxicologiques, et principalement la banque de toxicologie expérimentale, ont été effectivement utilisées pendant près de trois ans mais principalement par les services de l'INSERM. Il n'y a donc pas eu de commercialisation du produit, bien qu'au départ du projet, les utilisateurs potentiels ne manquaient pas, ni de rentabilisation de l'exploitation, bien qu'il existât des tarifs de question. Faut-il trouver une explication à ce fait ?

Que la demande d'informations toxicologiques ait disparu, reste totalement improbable. Que le système n'ait pas tenu compte des impératifs initiaux, nous n'oserions en douter, en tout état de cause. Peut-être la demande a-t-elle été stimulée par la distribution des questionnaires, peut-être un intérêt certain a-t-il été mis en lumière mais l'absence d'appui réel apporté par les demandeurs dans la conception même de la banque, en a restreint la vocation et la promotion qui devait en être faite.

D'autre part, la mise sur le marché de services d'informations de ce type n'est pas encore suffisamment courante, du moins en Europe, pour prétendre à un autofinancement



réel, dans un domaine finalement assez particulier.

## G. Rentabilité. Résultats.

### §1. Evaluation du coût.

#### - Personnel :

\* Informaticien : 4 années-hommes pour la mise au point des programmes de chargement et d'interrogation.

\* Scientifique : 1 année-homme pour la mise au point du code utilisé (en collaboration avec les informaticiens).

4 années-hommes (pharmacien, biologiste) pour la saisie des données.

#### - Heures-machines : ???

### §2. Evaluation des avantages.

Il faudrait opposer ces coûts à ceux qui sont économisés grâce à l'utilisation du système :

- temps passé par les chercheurs à la recherche de documents et à leur analyse,
- coût des expériences toxicologiques (entre 50.000 et 100.000 \$ pour une demande d'autorisation de mise sur le marché),
- les intérêts secondaires non négligeables (cfr §8).

### §3. Performance offertes par le système.

Les programmes SINBAD II permettent l'accès direct et isotrope aux données en limitant - les coûts,

- la place mémoire nécessaire,
- le temps d'accès.

Le mode de stockage adopté et les économies d'écriture permettent de gagner plus de 50% de la place mémoire normalement requise (chiffre valable pour le nombre de données effectivement enregistrées).

Une étude a laissé entrevoir que ce pourcentage ne peut que

s'accroître avec le nombre des entrées.

Le temps de réponse est quasi-indépendant du nombre de critères de sélection introduits dans la question et du volume de la réponse . Il est de l'ordre de la seconde.

En outre, grâce aussi à une minimisation du volume des transferts de données en mémoire, les coûts de traitement globaux ont été sensiblement réduits par rapport à un S.G.B.D. général.

#### §4. Quelques chiffres (B.D.D.).

- La B.D.D. contient 400 articles décrivant 1683 sous-articles. Les données se répartissent en 1350 colonnes (classes) regroupées en 36 grilles.
- Près de 10000 valeurs distinctes garnissent l'ensemble de ces colonnes.
- Le BLOC-LIGNES comporte 39250 lignes après compression.
- L'ensemble des fichiers internes occupe 578815 mots (le REFCLAIR : 80640 mots, le BLOC-LIGNES 161280 mots).
- Le chargement complet (sans ennui) se fait en 50 minutes.

#### H. Intérêts secondaires.

Ce sont des intérêts généralement valables pour les banques de données médicales :

- Validation des données existantes :
  - aide à la mise au point du protocole expérimental,
  - étude de corrélations entre différents facteurs,
  - étude de l'influence de nouveaux facteurs,
  - ...etc.
- Possibilité de comparer les résultats d'une expérience nouvelle aux données antérieures présentes dans la banque, en vue d'aider à l'interprétation des résultats de cette expérience.
- Découverte de nouvelles voies de recherche ou orientation des recherches, en montrant des lacunes à combler, en suggérant des hypothèses à tester.



- Mise au point d'une méthode d'analyse de résultats (et de présentation des résultats).

## I. Critique globale du système.

Nous voudrions ici mettre en lumière, non plus les performances offertes par le logiciel SINBAD, mais souligner quelques points faibles.

Le problème de la mise à jour a été tout simplement escamoté par les concepteurs du système. S'il est vrai qu'elle ne constitue pas une exigence et que par ailleurs sa fonction se limite à l'insertion de nouveaux éléments, le compromis adopté à l'avantage exclusif de l'interrogation de la banque de données nous paraît un peu excessif. Les insertions qui ne peuvent se faire qu'en mode différé doivent être groupées au maximum en sessions, de sorte que certains fichiers ne soient pas constamment recopiés.

Le contrôle des données, tel qu'il est opéré à l'heure actuelle, n'est pas suffisamment élaboré. Chaque valeur entrée est bien évidemment validée, quelques contrôles croisés sont également réalisés, mais il faudrait que le filtrage global soit beaucoup plus serré.

A ce point de vue, les scientifiques pourraient apporter d'appréciables conseils.

Une fois de plus, il nous faut déplorer l'absence d'un véritable dossier de conception et de programmation. Le système a été imaginé, testé et amélioré par les deux mêmes personnes. Les modifications périodiquement apportées font que seules ces deux personnes connaissent à peu près chacun de ses composants. Leur système restera donc toujours leur propriété, incontrôlable.

Au niveau de l'interrogation, quelques améliorations sont envisageables. Nous pensons à une présentation plus synthétique (tableau, graphe,...) des réponses trop longues ou trop vastes, ou encore à un module de réorientation auto-

matique des questions.

Il faut tout de même reconnaître que le nombre d'articles effectivement enregistrés (300) ne constitue pas un échantillon suffisant. Il en faudrait bien davantage pour pouvoir confirmer les performances annoncées.

Il est probable que s'accroisse, avec le nombre d'articles enregistrés, le taux de compression du BLOC-LIGNES mais, d'un autre côté, le nombre des pointeurs et des segments à gérer augmenterait de façon non prévisible (REFCOD et REFCLAIR).

Tout comme la banque des C.A.P. il s'agit plutôt d'une expérience pilote qui a mis en lumière des besoins et des solutions. Le financement et la décision de pareils projets relèvent d'autres responsables.



Jour de la semaine						
LU	MA	ME	JE	VE	SA	DI
16	17	18	19	20	21	22

Date		
J	M	Heure

Langue						
inc	F	N	All	Angl	autre	
31	32	33	34	35	36	

GENERALITES : 9

n° fiche mère

appel par 39 40 41 42 43 44

tél. lettre radio télex vis. CAP autre

appel urgent 23 27 29

oui non 45 46

code int. 52

Code médecin

DEMANDEUR :

Nom : . . . . .

Adresse : . . . . .

N° postal 22

N° téléphone : . . . . .

inconnu 26 anonymat 27 région 28

famille public-voisins intoxicqué médecin non précisé spécialiste généraliste méd. du travail vétérinaire pharmacien infirmier paramédicaux

41 43 45 46

hôpital collectif entreprise secours public C.A.P. autre

41 43 45 46

DEMANDE : 9

analyse information thérapeutique 4 procurer médicament bibliographie éducation sanitaire enseignement composition prévention renseignement ne conc. pas le CAP cas médical non toxique identification de produit divers épidémiologie rens. Centre

11 13 15 17 19 21

L'INTOXIQUE :

Nom : . . . . . Profession : . . . . . Nationalité : . . . . .

Age : 60

masculin 57 féminin 58 sexe inc. 59

adulte 52 enfant 54 prématuré 55 foetus 56 animal 50

Profess. Nat. 47 50

Poids kg. 68 Obèse 69 Maigre 69

ANTECEDENTS :

9 durée de grossesse (en mois) grossesse durée inconnue

11 allergie-idiosyncrasie diabète anomalie génétique

14 maladie psychiatrique (dépression exclue) dépression

16 toxicomanie éthylisme nuisance tox. prof. et pollution

19 digestifs neurologiques endocrinologiques

21

CIRCONSTANCES : Date et heure

34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55

EXPOSITION : 44 unique 45 à répétition 46 chronique

47 brève 48 prolongée

INTOXICATION VOLONTAIRE : 49 conduite suicidaire 50 criminelle 51 toxicomanie 52 abstinence (syndrome) 53 avortement 54 autre 55 professionnelle

INTOXICATION ACCIDENTELLE : 55

PRODUITS : C.C.7

Dénomination

dose exacte con-nue	quantités					voie	type	com-position Incon.	
	0	1	2	3	4				5
31	32	33	34	35	36	37	38	40	42

--	--	--	--	--	--	--	--	--	--



CC3

23	ténaux	57	accident thérapeutique
25	cardio-vasculaires		erreur thérapeut. (date pérempt.,
27	pulmonaires	59	voie, produit, surdosage)
	hématologiques		autre . . . . .
	dermatologiques		Intoxication indéterminée
	infectieux		
29	traitement médicamenteux		
	par . . . . .		
30	impregnation toxique		
31	autres caractéristiques		
32	récidive		
	débilité mentale		

LIEU : 61 domicile  
63 voie publique  
65 hôpitaux - cliniques  
67 école, crèche, collectivité  
69 laboratoire  
nature (campagne)  
lieu de travail  
autre . . . . .

APPEL : 71 au producteur  
73 à autre C.A.P.  
74 autre . . . . .  
74 intervention d'un autre organisme . . . . .

SYMPTOMATOLOGIE : 9 Nulle 10 Présente 11 Inconnue  
TRAITEMENT (C. I. n./respectées) : 12 avant appel 13 après conseil C.A.P.

TRANSPORT : 14 voiture ou ambulance  
15 avec surveillance méd. spécialisée  
16 avion ou hélicoptère

LIEU DE TRAITEMENT : 17 sur place  
18 intervention médicale  
19 ou vétérinaire  
20 hôpital  
21 transfert  
22 autre

proposé réel

avant appel

TRAITEMENT (résumé)

avant con- effec-  
appel seillé tué

27 évacuateur  
30 épurateur  
33 neutralisation-absorption  
36 antidotes-chélateurs  
39 symptomatique  
42 arrêt trait. motivant appel  
45 abstention thérapeutique  
48 traitement inconnu ou non défini

Mesures préventives 51

Détails en clair :

BILAN :

a) demande de nouvelles : non 9 par téléphone 10 par lettre 11 obtenues : oui 12 non 13

b) analyse toxicologique : conseillée 14 faite à A. S. B. L. 15 U. L. B. 16 V. U. B. 17 Gand 18 Liège 19 Louvain 20 autre 21

c) médicament envoyé : 32 n° analyse 22 23 24 25 26 27 28 29 30 31 non déterminable 37

d) relation : certaine 33 probable 34 douteuse 35 nulle 36

e) validité du cas : bonne 38 satisfaisante 39 nulle 40

f) évolution : guérison 42 séquelles 43 décès 44 inconnu 45

g) cas très intéressant : 46

Fiche complémentaire remplie 41

CC5

CC4



Annexe 6.

LEGENDE \* = NOMBRE DE DOSSIERS  
REPONANT ENCORE AUX CRITERES  
EXIGES

T = TEMPS TOTAL  
i = CRITERE i  
M = MOYENNE (par  
FONDERATION)

ORGANISATION TYPE-1				ORGANISATION TYPE-2				ORGANISATION TYPE-3				
	H <sub>a</sub>	H <sub>b</sub>	H <sub>c</sub>	H <sub>a</sub>	H <sub>b</sub>	H <sub>c</sub>	H <sub>a</sub>	H <sub>b</sub>	H <sub>c</sub>	H <sub>a</sub>	H <sub>b</sub>	H <sub>c</sub>
Q <sub>1</sub> →	T	16.016	16.016	16.016	T	16.016	16.016	16.016		6.194	6.194	6.194
Q <sub>2</sub>	1	26.180	26.180	26.180					4.313	4.313	4.313	
	*	660	1980	11000				660	1980		11000	
	2	8.734	8.773	8.773	IL SUFFIT DE NE TRANSFERER QU'UNE SEULE FOIS CHACUNE DES PAGES DU FICHIER PRINCIPAL T-APPEL			3.316	3.316	3.316		
	*	7	60	1100		7	60		1100			
	3	0.245	0.956	10.251		3	0.157	0.846	3.316			
	*	1	15	110		1	1	15	110			
→	T	35.159	35.909	45.204	T	66.100	66.100	66.100	7.786	8.475	10.945	
Q <sub>3</sub>	1	16.016	16.016	16.016	1	16.016	16.016	16.016	6.194	6.194	6.194	
	*	6	17	11000	*	6	17	11000	6	17	11000	
	2	0.210	0.441	12.4532	2	0.210	0.595	66.100	3	0.166	0.309	5.236
	*	2	13	1100	*	2	13	1100	1	2	1100	
	3	0.070	0.421	6.8403	3	0.070	0.455	14.630	2	0.035	0.048	0.672
	*	1	1	110	*	1	1	110	1	1	110	
→	T	16.296	16.878	35.309								

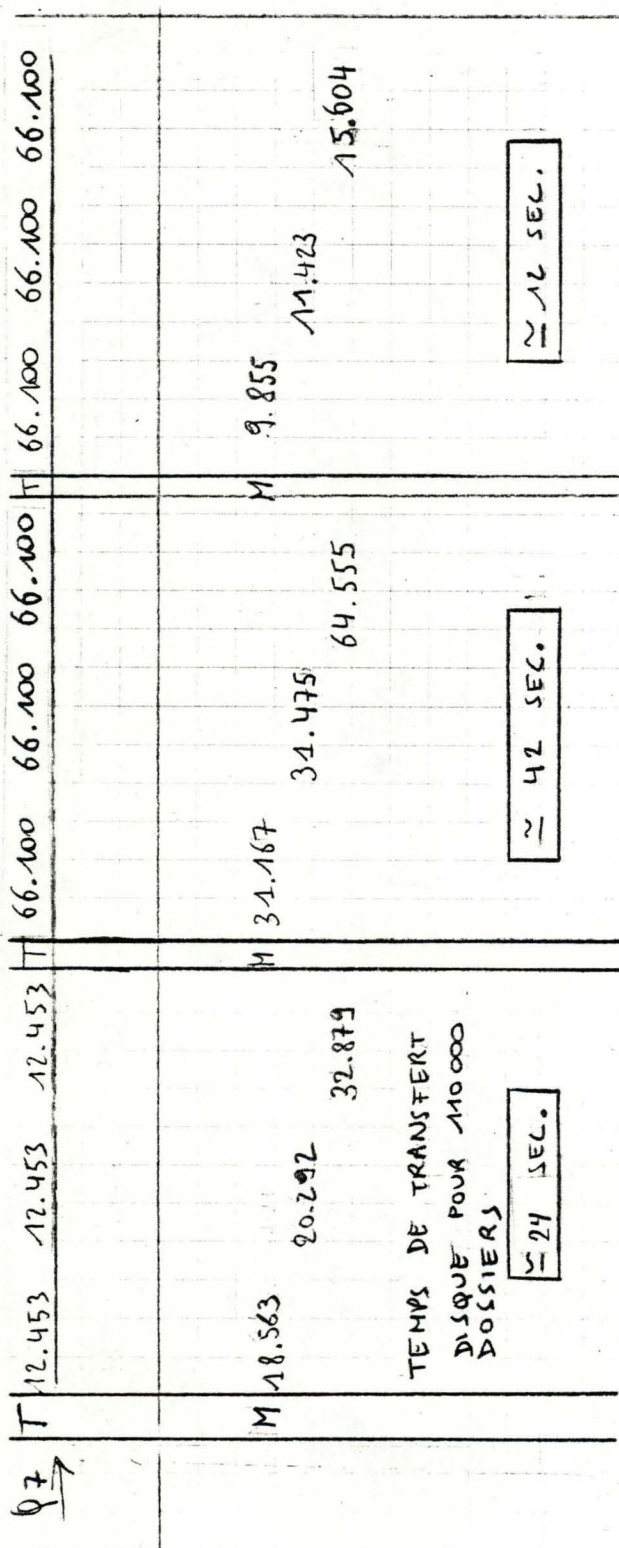
IL FAUT Y AJOUTER LE TEMPS D'ACCES A T-TRAITEMENT RENDU NECESSAIRE PAR LE FAIT QUE T-TRAITEMENT N'EST PAS DIRECTEMENT ACCESSIBLE PAR LE NO-DOSSIER +0.070 +0.389 +12.453				IL FAUT Y AJOUTER LE TEMPS D'ACCES A T-DOSSIER RENDU NECESSAIRE PAR LE FAIT QUE T-COMPLEM N'EST PAS DIRECTEMENT ACCESSIBLE PAR LE NO-DOSSIER. +0.070 +0.455 +15.400							
I	16.366	17.267	47.762	T	16.366	17.521	112.146	T	6.395	6.551	12.102
1	16.016	16.016	16.016	1	16.016	16.016	16.016	1	6.194	6.194	6.194
*	6	17	11000 *	6	17	11000	*	6	17	11000	
2	0.210	0.375	8.773 1/3	0.210	0.595	66.100	4	0.188	0.277	6.610	
*	1	1	1100 *	1	1	110	*	1	1	1100	
3	0.035	0.035	12.453 1/5	0.035	0.035	2.810	2	0.035	0.035	3.316	
*	1	1	110 *	1	1	2	*	1	1	110	
I	+0.035	+0.035	+1.650	+0.035	+0.035	+2.530	5	0.035	0.035	1.518	
4	0.035	0.035	1.854	-	-	-	*	1	1	11	
*	1	1	11	-	-	-	3	0.035	0.035	0.165	
5	0.035	0.035	0.395	-	-	-	*	1	1	2	
*	1	1	2	-	-	-	-	-	-	-	
→	16.366	16.531	41.141	T	16.296	16.681	87.456	T	6.487	6.576	17.803



IL FAUT Y AJOUTER LE TEMPS D'ACCES A T-TRAITEMENT RENDU NECESSAIRE PAR LE FAIT QUE T-COMPLEM N'EST PAS DIRECTEMENT ACCESSIBLE PAR LE NO-BOSSIER.		IL FAUT Y AJOUTER LE TEMPS D'ACCES A T-BOSSIER RENDU NECESSAIRE PAR LE FAIT QUE T-COMPLEM N'EST PAS DIRECTEMENT ACCESSIBLE PAR LE NO-BOSSIER.	
+0.070 + 0.389 + 12.453		+0.070 + 0.455 + 15.400	
I	16.366	17.267	47.762
T	16.366	17.267	47.762
1	16.016	16.016	16.016
*	6	17	11000
2	0.210	0.375	8.773
*	1	1	1100
3	0.035	0.035	12.453
*	1	1	110
I	+0.035	+0.035	+1.650
4	0.035	0.035	1.854
*	1	1	11
5	0.035	0.035	0.395
*	1	1	2
T	16.366	16.531	41.141

	IL FAUT Y AJOUTER LE TEMPS D'ACCES A T-DOSSIER RENDU NECESSAIRE PAR LE FAIT QUE T-COMPLEM N'EST PAS DIRECTEMENT ACCESSIBLE PAR LE NO-DOSSIER.				IL FAUT Y AJOUTER LE TEMPS D'ACCES A T-DOSSIER RENDU NECESSAIRE PAR LE FAIT QUE T-COMPLEM N'EST PAS DIRECTEMENT ACCESSIBLE PAR LE NO-DOSSIER.			
	16.366	17.267	47.762	16.366	17.521	16.016	16.016	16.016
I	16.366	17.267	47.762	16.366	17.521	16.016	16.016	16.016
T	16.366	17.267	47.762	16.366	17.521	16.016	16.016	16.016
1	16.016	16.016	16.016	16.016	16.016	16.016	16.016	16.016
*	6	17	11000	6	17	11000	6	11000
2	0.210	0.375	8.773	0.210	0.595	66.100	0.210	66.100
*	1	1	1100	1	1	110	1	1100
3	0.035	0.035	12.453	0.035	0.035	2.810	0.035	3.316
*	1	1	110	1	1	2	1	110
I	16.035	16.035	16.035	16.035	16.035	16.035	16.035	16.035
4	0.035	0.035	1.854	0.035	0.035	2.530	0.035	1.518
*	1	1	11	1	1	1	1	11
5	0.035	0.035	0.395	0.035	0.035	0.035	0.035	0.165
*	1	1	2	1	1	1	1	2
T	16.366	16.531	41.141	16.296	16.681	87.456	16.296	17.803





Φ5	Φ5 REMARQUE FAITE EN Q2				Φ6 REMARQUE FAITE EN Q2			
	1	8.773	8.773	8.773	1	8.773	8.773	8.773
*	*	1100	3300	11000	*	1100	3300	11000
2	2	10.251	10.251	10.251	2	10.251	10.251	10.251
*	*	88	825	1100	*	88	825	1100
3	3	1.452	11.033	14.608	3	2.244	11.825	15.400
*	*	2	50	110	*	2	50	110
→	T	20.476	30.057	33.632	T	66.100	66.100	66.100
Φ6	CFR. Q2 POUR 1/2/3				CFR. Q2 POUR 1/2/3			
	1/2/3	35.159	35.909	45.204	1/2/3	7.786	8.475	10.945
*	*	1	15	110	*	1	15	110
4	4	0.035	0.035	1.650	4	0.035	0.305	0.760
*	*	1	1	11	*	1	1	11
5	5	0.035	0.035	0.385	5	0.035	0.035	0.385
*	*	1	1	2	*	1	1	2
6	6	0.035	0.035	0.070	6	0.035	0.035	0.070
*	*	1	1	1	*	1	1	1
→	T	35.264	36.04	47.309	T	7.891	8.850	12.160



Répertoire des grilles (Banque des C.A.P.)

1. intoxiqué
2. intoxications collectives
3. appel demandeur
4. antécédents/terrain
5. produit absorbé
6. mélange de produits
7. bilan
8. troubles hépato-digestifs
9. signes cutanés et généraux
10. signes endocriniens
11. signes neurologiques et musculaires
12. signes neurosensoriels
13. signes ostéo-articulaires
14. signes respiratoires
15. signes cardiovasculaires
16. troubles génito-urinaires
17. signes psychiques
18. examens paracliniques
19. examens toxicologiques
20. examens biologiques
21. traitement non symptomatique
22. traitement symptomatique

Répertoire des grilles (Banque B.D.S.)

1. protocole expérimental
2. hématologie
3. signes hépato-digestifs
4. signes cutanés et généraux
5. signes respiratoires
6. signes cardiovasculaires
7. signes psychiques
8. signes neurologiques et neuromusculaires
9. signes concernant l'oeil, le nez, l'oreille
10. signes endocriniens
11. signes ostéo-articulaires
12. signes génito-urinaires
13. cancérogénicité et tératogénicité
14. examens biologiques
15. traitement

16. distribution

17. excrétion

18. fertilité, anomalies de la reproduction

Répertoire des grilles (Banque B.D.D.)

- 1 Définition des produits
- 2 Définition des mélanges de produits
- 3 Conditions d'administration des produits
- 4 Sujets d'expérience
- 5 Traitements associés
- 6 Conditions de vie des animaux d'expérience
- 7 Doses léthales
- 8 Mortalité
- 9 Examens hématologiques 1
- 10 Examens hématologiques 2
- 11 Formule leucocytaire
- 12 Biochimie
- 13 Histopathologie
- 14 Examens respiratoires
- 15 Signes respiratoires
- 16 Signes génito-urinaires
- 17 Examens macroscopiques 1
- 18 Examens macroscopiques 2: Poids des sujets
- 19 Examens macroscopiques 3: Poids des organes
- 20 Électroencephalogramme
- 21 Électromyographie— excitabilité neuromusculaire
- 22 Signes neurologiques et neuromusculaires
- 23 Signes hépato-digestifs
- 24 Signes cutanés et généraux
- 25 Pharmacocinétique
- 26 Cancérogénicité
- 27 Examens cardiovasculaires— hémodynamique
- 28 Électrocardiogramme
- 29 Fréquence ou rythme cardiaque
- 30 Signes cardiovasculaires
- 31 Médullogramme
- 32 Anesthésie (avant examen)
- 33 Tolérance locale
- 34 Examens divers
- 35 Fertilité
- 36 Anomalies morphologiques de la reproduction
- 37 Bilan



BUMP



0 0 3 6 2 0 7 3 1

\*FM B16/1979/20

